

**Identifying small molecule binding ligands for the pup-ligase (PafA) of
*Mycobacterium tuberculosis***

by

SANDILE M. MTHEMBU

Submitted in full fulfilment of the academic requirements of

Master of Science

in Biochemistry

School of Life Science

College of Agriculture, Engineering and Science

University of KwaZulu-Natal

Pietermaritzburg

South Africa

10 February 2022

Supervised by Prof. Raymond Hewer and Dr. Alexandre Delport

PREFACE

The research contained in this dissertation was completed by the candidate while based in the Discipline of Biochemistry, School of Life Science of the College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Pietermaritzburg campus, South Africa. The research was financially supported by the National Research Foundation.

The contents of this work have not been submitted in any form to another university and, except where the work of others is acknowledged in the text, the results reported are due to investigations by the candidate.



Signed: Prof. R HEWER

Date: 10 February 2022



Signed: Dr. A DELPORT

Date: 10 February 2022

DECLARATION: PLAGIARISM

Note that two declaration sections are required if there are papers emanating from the dissertation/thesis. The first (obligatory) declaration concerns plagiarism and the second declaration specifies your role in the published papers.

I, Sandile Menzi Mthembu, declare that:

(i) the research reported in this dissertation, except where otherwise indicated or acknowledged, is my original work;

(ii) this dissertation has not been submitted in full or in part for any degree or examination to any other university;

(iii) this dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) this dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced;

(v) where I have used material for which publications followed, I have indicated in detail my role in the work;

(vi) this dissertation is primarily a collection of material, prepared by myself, published as journal articles or presented as a poster and oral presentations at conferences. In some cases, additional material has been included;

(vii) this dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation and in the References sections.



Signed: SANDILE MENZI MTHEMBU

Date: 10 February 2022

ABSTRACT

The rapid emergence of resistant TB strains (*Mycobacterium tuberculosis* (*Mtb*)) renders traditional treatment options ineffective and necessitates the generation of novel anti-TB drugs that possess innovative modes of action. The pup-ligase (PafA) of *Mtb* that solely mediates protein proteasomal removal via the pupylation cascade has recently been identified as a suitable target for TB drug development. A novel approach would be to recruit proteolysis targeting chimeras (PROTACs) technology as an alternative anti-TB treatment option by developing PROTAC-like molecules capable of recruiting the pupylation cascade. Therefore, the identification of novel PafA small-molecule binding ligands is an essential first step to establish possible new TB therapies. To this effect, PafA recombinant expression was successfully optimised in *E. coli* cells at 20°C for 20 h, where a 50-kDa protein was observed by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Moreover, the identity of the protein was confirmed via immunoblotting with anti-His antibodies and PafA subsequently purified via immobilized metal affinity chromatography (IMAC) to high purity. A thermal shift assay (TSA) of PafA against 48 small-molecule compounds from a chemical library pre-screened for non-specific inhibition activity was conducted. Seven Hit compounds were detected significantly binding PafA ($P < 0.05$), all inducing a > 5 °C increase of PafA melting temperature (T_m) upon binding. Future research on these novel PafA binding ligands will be to ascertain whether they possess inhibitor qualities. Additionally, they will be used in the synthesis of heterobifunctional molecules to create the first PROTAC-like molecules for targeted proteasomal degradation of essential *Mtb* proteins – a novel type of anti-TB drug.

ACKNOWLEDGMENTS

I would like to express sincere gratitude to my supervisors Prof. R Hewer and Dr A Delport for their guidance, deep knowledge, and brilliant advice which gave me the confidence to carry out this research project and face all its challenges head on. I am grateful for all your efforts in making the research project a reality.

I would also like to thank Dr C. Pillay for providing funding- through the NRF- without which the project would not have been a reality.

Much thanks to Mark, and the rest of my colleagues in lab 46 for their continued support and unyielding encouragement.

To my school colleagues, Nomfundo, Thokozani, Nomacusi, Sinenhlanhla, Thandolwethu, Sbahle, and Bongumusa. I am grateful for the support, input and motivation.

I am grateful to the National Research Foundation (NRF) for funding this project and the University of KwaZulu-Natal for providing laboratory facilities and equipment to facilitate this research.

Lastly, I would like to thank two people to whom my studies would not be possible if not for their constant love and support, my mom Mrs B.R Gwija and dad Mr V Gwija

TABLE OF CONTENTS

| | <u>Page</u> |
|--|-------------|
| PREFACE | ii |
| DECLARATION: PLAGIARISM | iii |
| ABSTRACT | iv |
| ACKNOWLEDGMENTS | v |
| TABLE OF CONTENTS | vi |
| ABBREVIATIONS | viii |
| LIST OF TABLES | x |
| LIST OF FIGURES | xi |
| CHAPTER 1: LITERATURE REVIEW | 1 |
| 1.1 Introduction | 1 |
| 1.2 Protein degradation cascade systems | 3 |
| 1.2.1 The ubiquitin proteasomal degradation system | 3 |
| 1.2.2. Pupylation: ubiquitin-like protein proteasomal degradation system | 4 |
| 1.3. Proteinopathy therapy | 5 |
| 1.3.1 Induced protein degradation in Eukaryotes | 6 |
| 1.3.1.1 Specific non-generic IAP-based protein erasers (SNIPERs) | 6 |
| 1.3.1.2 PROTACs | 7 |
| 1.3.1.2.1 Cereblon-engaging PROTACs | 9 |
| 1.3.1.2.2 VHL-engaging PROTACs | 10 |
| 1.3.2 Induced protein degradation in prokaryotes | 13 |
| 1.3.3 Pupylation and tuberculosis | 14 |
| 1.4 Study rationale | 16 |
| 1.5 Aims and Objectives | 16 |
| CHAPTER 2: MATERIALS AND METHODS | 17 |
| 2.1 Recombinant expression of PafA | 17 |
| 2.1.1 Vector recovery | 17 |
| 2.1.2 Competent cell preparation and transformation | 17 |
| 2.1.3 Polymerase Chain Reaction (PCR) and Sequencing of PafA and PupE | 18 |
| 2.1.4 Recombinant PafA expression | 21 |
| 2.1.5 SDS-PAGE and immunoblot analysis | 21 |

| | |
|---|----|
| <u>2.1.6 PafA solubilisation and purification</u> | 22 |
| <u>2.2 The detection of aggregating pan assay interfering compounds (PAINs) via a colorimetric enzyme assay</u> | 24 |
| <u>2.2.1 Enzyme assay</u> | 24 |
| <u>2.2.2 Non-specific inhibitor assay</u> | 24 |
| <u>2.3 Screening for potential PafA-binding ligands</u> | 25 |
| <u>2.3.1 PafA TSA optimization</u> | 26 |
| <u>2.3.2 Screening for PafA-binders</u> | 28 |
| <u>CHAPTER 3: RESULTS</u> | 29 |
| <u>3.1 Recombinant PafA expression, detection, and purification</u> | 29 |
| <u>3.1.1 Vector DNA isolation and characterisation</u> | 29 |
| <u>3.1.2 Recombinant PafA protein expression and immunoblot analysis</u> | 32 |
| <u>3.1.3 PafA solubilisation and purification</u> | 35 |
| <u>3.2 The detection of aggregating PAINs via a colorimetric enzyme assay</u> | 40 |
| <u>3.2.1 Enzyme assay</u> | 40 |
| <u>3.2.2 Non-specific inhibitor screening assay</u> | 41 |
| <u>3.3 Screening of small molecules for PafA-binding capability via a thermal shift assay (TSA)</u> | 43 |
| <u>3.3.1 PafA-binding ligand screening via the optimised PafA TSA</u> | 45 |
| <u>CHAPTER 4: DISCUSSION</u> | 49 |
| <u>4.1 PafA and PupE bacterial transformation</u> | 49 |
| <u>4.2 PafA recombinant expression and purification</u> | 50 |
| <u>4.3 PafA solubilisation and purification</u> | 52 |
| <u>4.4 Non-specific inhibitor assay</u> | 54 |
| <u>4.5 Screening for PafA-binding ligands via a TSA</u> | 56 |
| <u>REFERENCES</u> | 59 |
| <u>APPENDIX</u> | 68 |

ABBREVIATIONS

| Abbreviation | Meaning |
|----------------------|---|
| 1xEB | 1 x equilibration buffer |
| BacPROTAC | Bacterial PROTAC |
| BRD4 | Bromodomain 4 |
| Cas9 | CRISPR associated protein 9 |
| CHAMPs | Chaperone-mediated protein degradation |
| cIAP1 | Cellular IAP1 |
| CRBN | Cereblon |
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| CRL | Cullin-RING |
| CRL2 ^{VHL} | VHL-CRL2 complex |
| CRL4 ^{CRBN} | Cereblon-CRL4 complex |
| CUL4A | Cullin-4A |
| DDB1 | DNA-binding protein 1 |
| DMSO | Dimethyl sulfoxide |
| Dop | Deamidase of pup |
| E1 | Ubiquitin activating enzyme |
| E2 | Ubiquitin-conjugating enzymes |
| E3 | Ubiquitin-ligase |
| ECL | Enhanced chemiluminescence |
| EGCG | Epigallocatechin gallate |
| HRP | Horse radish peroxidase |
| IAP | Inhibitor of apoptosis |
| IKZF1 | IKAROS Family Zinc Finger 1 |
| IKZF3 | IKAROS Family Zinc Finger 3 |
| IMAC | Immobilised-metal affinity chromatography |
| IMiDs | Immunomodulatory drugs |
| K_i | Compound inhibitor constant |
| MDM2 | Mouse double minute 2 homolog |
| ME-BS | Bestatin methyl ester |
| Mpa | <i>Mycobacterium</i> proteasomal ATPase |
| <i>Mtb</i> | <i>Mycobacterium tuberculosis</i> |
| NRF | National Research Foundation |
| PafA | Pup-ligase |
| PAINs | Pan assay interfering compounds |
| PBS | Phosphate buffered saline (pH7.4) |
| PCR | Polymerase chain reaction |
| PLK1 | Polo-like kinase 1 |
| PROTACs | Proteolysis-targeting chimeras |
| Pup | Prokaryotic ubiquitin-like protein |

| | |
|------------------|---|
| PupE | Prokaryotic ubiquitin-like protein-Glutamate |
| RIPK1 | Receptor interacting protein kinase 1 |
| RNA | Ribonucleic Acid |
| RNAi | RNA interference |
| RNAP | DNA-dependent RNA polymerase |
| ROC1 | Regulator of cullins 1 |
| SAMP1 | Small archaeal modifier proteins 1 |
| SAMP2 | Small archaeal modifier proteins 2 |
| SDS | Sodium dodecyl-sulfate |
| SDS-PAGE | Sodium dodecyl-sulfate polyacrylamide gel electrophoresis |
| Sirt2 | Sirtuin2 |
| SNIPERs | IAP-based protein erasers |
| SOC | Super optimal cataboliser medium |
| TB | Tuberculosis |
| TBS | Tris-Buffered Saline |
| TBS _t | 0.001% (v/v) Tween-20 supplemented TBS buffer |
| TE | Tris-EDTA buffer (pH7.4) |
| T _m | Primer melting temperature |
| T _m | protein melting temperature |
| TMB | 3,3',5,5'-Tetramethylbenzidine |
| tmRNA | Transfer-messenger Ribonucleic acid |
| tRNA | Transfer Ribonucleic acid |
| TSA | Thermal shift assay |
| UbaA | SAMP-activating enzyme E1 |
| UP | Ubiquitin-proteasome cascade |
| VHL | Von Hippel Lindau |

LIST OF TABLES

| <u>Table</u> | <u>Title</u> | <u>Page</u> |
|---------------------|--|--------------------|
| Table 1.1 | Summary of Mtb mutated genes conferring resistant against first-line defence anti- <i>Mtb</i> drugs | 15 |
| Table 2.1 | Recombinant plasmid PCR primers. | 19 |
| Table 2.2 | Reaction setup for the control protein TSA | 26 |
| Table 2.3 | TSA Reaction setup for PafA-binding compounds..... | 28 |
| Table 3.1 | DiverSET CHEMBRIDGE compounds that significantly increased PafA T _m .. | 46 |
| Table A1 | Effect of DiverSET CHEMBRIDGE library compounds on PafA T _m | 70 |
| Table A2 | The portion of the DiverSET CHEMBRIDGE small molecule library available in our laboratory. | 71 |

LIST OF FIGURES

| <u>Figure</u> | <u>Title</u> | <u>Page</u> |
|----------------------|--|--------------------|
| Figure 1.1 | A schematic representation of the ubiquitin-proteasome cascade | 3 |
| Figure 1.2 | A schematic overview of the pupylation cascade | 4 |
| Figure 1.3 | PROTAC mediated protein degradation..... | 8 |
| Figure 1.4 | Phase I/II Arvinas clinical PROTACs | 12 |
| Figure 1.5 | <i>Mycobacterium</i> drug resistance mechanisms..... | 15 |
| Figure 2.1 | PCR reaction conditions for the amplification of (A) PafA and (B) Pup from the pET24b(+)-PafA and Pet24b(+)-PupEPafA constructs, respectively | 20 |
| Figure 2.2 | Thermal melt conditions for PafA thermal shift assays..... | 27 |
| Figure 3.1 | Agarose gel electrophoresis of pET24b(+)-PafA (A) and pET24b(+)-PupEPafA (B) recombinant <i>E. coli</i> cells..... | 30 |
| Figure 3.2 | Amplification by PCR of the PafA (A) and pup (B) genes from the pET24b(+)- PafA and pET24b(+)-PupEPafA plasmids, respectively | 31 |
| Figure 3.3 | Recombinant JM109 (DE3) and BL21(DE3) expression of PafA, from pET24b(+)- PafA, as analysed by SDS-PAGE and immunoblot | 32 |
| Figure 3.4 | Recombinant JM109 (DE3) <i>E. coli</i> expression of PafA, from pET24b(+)- PupEPafA, as analysed by SDS-PAGE and immunoblot..... | 34 |
| Figure 3.5 | The solubility of PafA when IPTG concentrations were varied (4 h, 37 °C). 35 | |
| Figure 3.6 | PafA solubility at low temperature expressions (20 h, 20 °C) | 36 |
| Figure 3.7 | SDS solubilisation of PafA from inclusion bodies..... | 37 |
| Figure 3.8 | PafA IMAC purification using TALON resin..... | 39 |

| | | |
|-------------|---|-----------|
| Figure 3.9 | The effect of HRP enzyme concentration on colourimetric development in the presence of TMB substrate | 40 |
| Figure 3.10 | The effect of non-specific on the colourimetric HRP enzyme assay | 42 |
| Figure 3.11 | Screening 48 compounds from the DiverSET CHEMBRIDGE library for non-specific inhibition | 43 |
| Figure 3.12 | Control thermal shift assay using the protein thermal shift starter kit | 44 |
| Figure 3.13 | Optimised PafA thermal shift assay | 45 |
| Figure 3.14 | Screening PafA for possible binding ligands using the TSA..... | 45 |
| Figure 3.15 | Molecular structure of the potential PafA binders..... | 47 |
| Figure A.1 | Vector construct size estimation | 68 |
| Figure A.2 | Recombinant PafA size estimation using known protein standards | 68 |
| Figure A.3 | The effect of DiverSET CHEMBRIDGE library compounds on PafA thermal melt reaction..... | 69 |
| Figure A.4 | The calculated change in T_m of PafA in the presence of the DiverSET CHEMBRIDGE compounds | 69 |

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

Protein homeostasis is a vital aspect of cellular health achieved via tightly regulated processes at the transcriptional, translational, and post-translational level. Post-translational protein modification is a well-established means for protein regulation and enhanced functional diversification that affects protein durability, activity, and subcellular localization (Burns and Darwin, 2010). In particular, modifications serving as quality control systems by flagging proteins for degradation have been the subject of extensive biological research (Rape, 2018; Deng *et al.*, 2020; Bhaduri and Merla, 2021).

In eukaryotes, ubiquitination is the best characterised and most conserved degradation system that exploits a small adaptor protein. In this system ~ 8-kDa ubiquitin, is repeatedly conjugated onto target proteins to form a polyubiquitin chain that flags the protein for energy dependent 26S proteasomal degradation (George *et al.*, 2018). The ubiquitin-proteasome (UP) cascade is explored in drug discovery processes to alleviate protein-associated diseases (Huang and Dixit, 2016). One such exploration has been the design and development of heterobifunctional molecules that recruit the UP system for selective degradation of a bound target protein (Bondeson and Crews, 2017; Bond and Crews, 2021).

Macromolecular protein modifiers for targeted proteasomal degradation, a regulatory system well described for eukaryotes, has been reported in prokaryotes as well (Pearce *et al.*, 2008). Observed in select members of the actinobacteria family, pupylation was the first to be described where a small **prokaryotic ubiquitin-like protein** “pup” is conjugated to target proteins to signal proteasomal degradation. Though functional analogues, pup and ubiquitin display distinct modification pathways of their targets. The discovery and continual delineation of the pupylation system has sparked interest in the research of other ubiquitin-like proteins in prokaryotes, namely the small archaeal modifier proteins SAMP1 and SAMP2 (SAMPs) (Maupin-Furlow, 2014).

SAMPs are widespread in archaea, conjugated to lysine residues of target proteins using the SAMP-activating enzyme E1 (UbaA), similar to ubiquitin’s conjugation to protein targets by

the ubiquitin activating enzyme (E1) (Humbard *et al.*, 2010). It is currently unknown, however, what the function of the cascade is but due to the high similarities it has with the ubiquitination system, it has been predicted to be involved in proteasome-mediated protein degradation. This hypothesis is supported by (1) the homology of the UbaA and E1 enzyme of ubiquitination, implying a close functional relationship, (2) the presence of proteasomes in archaea suggesting mediated degradation, and (3) both SAMPylation and ubiquitination targeting lysine residues of proteins. The SAMPylation system also displays similarities to the pupylation system found in select actinobacteria i.e., both systems use a single enzyme to affect small protein conjugation to lysine residues of target proteins. As further evidence, independent nitrogen growth studies done by Elharar *et al* and Humbard *et al* on *Mtb* and archaea, respectively, returned a strikingly similar response (Elharar *et al*, 2014; Humbard *et al*, 2010). In the absence of nitrogen, Elharar and coworkers reported elevated levels of pupylated proteins in *Mycobacterium smegmatis*, suggesting cytosolic protein turnover via the Mpa-20S proteasome for amino acid recycle (Elharar *et al.*, 2014). In the same study, a pupylation deficient *Mycobacterium smegmatis* strain with the genes for Pup and 20S proteasome knocked out, showed considerable defected growth under the same nitrogen deficient conditions compared to the control pupylating strain (Elharar *et al.*, 2014). Similarly, Humbard and colleagues observed significantly increased SAMPylated proteins in the archaeon *Haloferax volcanii* under nitrogen limited conditions, suggesting a similar proteasomal recycling system facilitated by the SAMPylation cascade (Humbard *et al*, 2010)

Prokaryotes also employ a highly conserved degradation-inducing tagging system involving an 11-amino acid peptide (SsrA) via a specialized tRNA molecule (tmRNA) (Hayes and Keiler, 2010). The tmRNA, recruited in the case of a defective translational product, is a bifunctional RNA molecule possessing both mRNA and tRNA activity to mediate the addition of the SsrA peptide to the C-terminus of nascent polypeptide chains. The SsrA flags the defective translated product for C-terminal degradation by specific endogenous proteases such as ClpA (Burns and Darwin, 2010).

The study and characterisation of these regulatory systems play a pivotal role in understanding and mediating various disease interventions. As a foundation for the ensuing study, this review discusses proteasomal degradation cascades, focusing on the ubiquitination and pupylation

systems. Moreover, the development and future prospect of specific drug mediated therapy that recruits degradative cascades is also examined.

1.2 Protein degradation cascade systems

1.2.1. Protein turnover in eukaryotes: The ubiquitin proteasomal degradation system

In cells polypeptides are processed, folded, modified, and shuttled to their target sites to carry out a specific function. Since a large proportion of synthesised proteins are required transiently, refined degradation systems exist to regulate protein turnover. The UP cascade stands as the largest (~1% of cellular biomass) and best conserved protein degradation system in eukarya (Bondeson and Crews, 2017). This system uses a cascade of enzymes i.e., E1, Ubiquitin-conjugating enzymes (E2), and ubiquitin-ligases (E3-ligase), which collectively mediates the repeated conjugation of ubiquitin to a target protein (Figure 1.1) (Kelley, 2018). Ubiquitin is conjugated onto a lysine residue of a target protein via an isopeptide bond, a process that is repeated several times to yield a polyubiquitin chain due to ubiquitin itself comprising of a number of lysine residues. This causes the protein to be recognized and unfolded by the capped ATPase of the proteasome, whilst being directed inside the 26S proteolytic core for degradation.

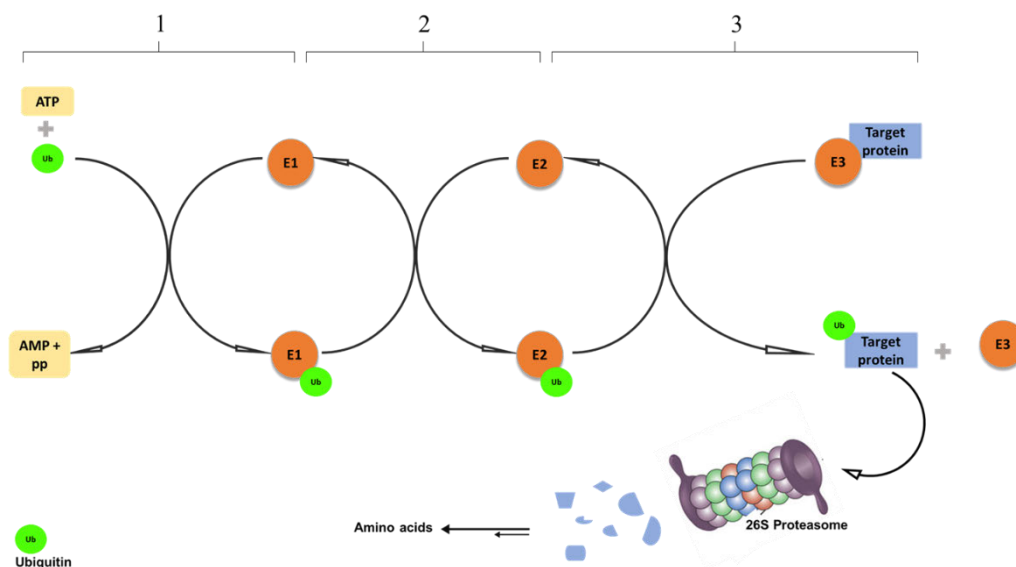


Figure 1.1. A schematic representation of the ubiquitin-proteasome cascade. Ubiquitin (in green) is activated for conjugation by the E1 enzyme using ATP (1), which subsequently transfers it to the E2 ubiquitin conjugating enzyme (2). The E3 ligase enzyme recruits the target protein and interacts with the ubiquitinated E2 ubiquitin conjugating enzyme, resulting in the target proteins ubiquitination (3). Ubiquitinated proteins are fated for degradation by the 26S proteasome (Adapted from Kelly, 2018).

1.2.2 Protein turnover in prokaryotes: the pupylation system (ubiquitin-like proteasomal protein degradation system)

Macromolecular protein modification in prokaryotes was first described by Pierce and colleagues who reported the discovery of a small ubiquitin-like protein “pup” that could specifically conjugate to proteasomal targets in *Mtb* (Pearce *et al.*, 2008). This pup-proteasome cascade is a post-translational protein modification system limited to members of the phylum actinobacteria, a group of prokaryotic organisms that are typically symbionts with eukaryotes (Cui *et al.*, 2021). Perhaps, one of the most fascinating concepts about this taxonomic group is the expression of compartmentalised proteases, proteosomes, typically only found in eukaryotes and archaea (Barandun *et al.*, 2012). Pup conjugation to a target protein is facilitated by PafA, and follows two steps. As seen from Figure 1.2, prior to conjugation pup is activated via deamidation of its C-terminal glutamine residue to glutamate using deamidase of pup (Dop). The activated pup is then recruited by PafA (blue) and conjugated to a specific lysine residue of a target protein (grey).

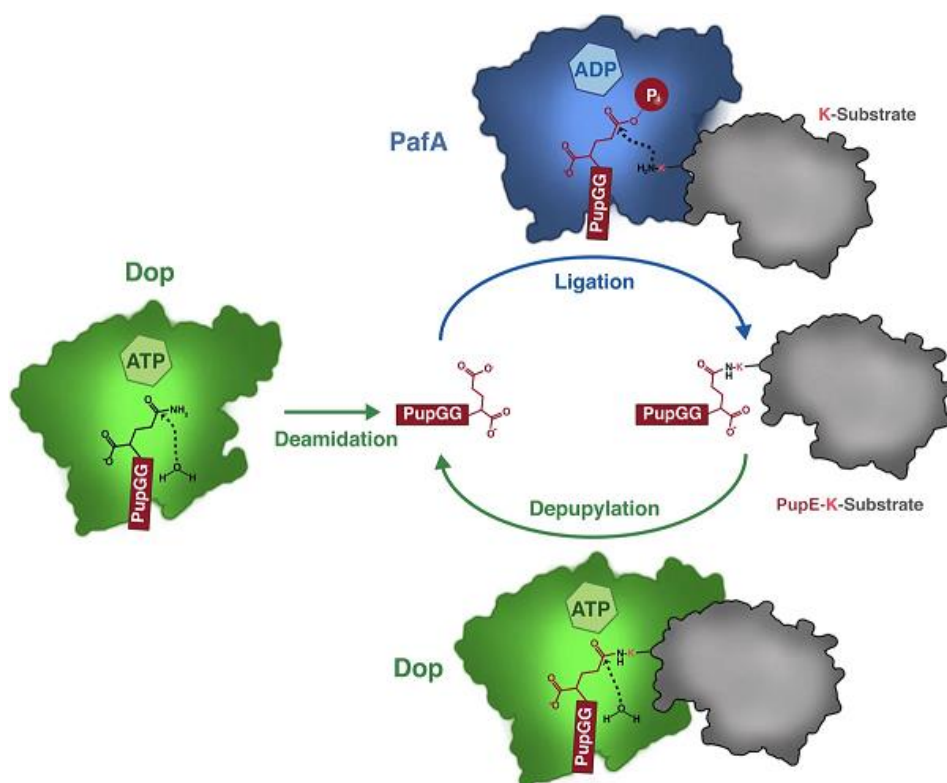


Figure 1.2. A schematic overview of the pupylation cascade. Pup is initially activated via Dop deamidation of its C-terminal Gln, and subsequently recruited by PafA which ligates the deamidated pup to a specific lysine (Lys) residue of target substrates. Pupylated proteins are flagged products for proteasomal degradation (Taken from Barandun *et al.*, 2012).

There has been extensive research focusing on the characterisation of the pupylation pathway. In their pioneering work, Pierce and colleagues not only established and demonstrated that enzymes Dop and PafA were responsible for pup's conjugation to a protein substrate, but also offered insight as to how the pupylated products are delivered to the proteasome for degradation (Pierce *et al.*, 2008). Through various experiments, Pierce and coworkers observed that PafA deficiency in *Mtb* mutants resulted in the absence of pupylated substrates, but accumulated pup-free established targets (Pierce *et al.*, 2008). This observation not only demonstrated PafA as the conjugating factor for pupylation but also suggested the enzyme to be the sole ligase responsible for pupylating substrates in *Mtb*. This vastly contrasts with the ~600 different E3 ligases of the UP system found in eukaryotes. The group additionally discovered that pup associated with *Mycobacterium* proteasomal ATPase (Mpa), a protein sharing homology with eukaryotic proteasome ATPases that recognise, prepares, and directs substrates into the 20S proteasome core for degradation. In Mpa deficient *Mtb* mutants, saturated levels of pupylated substrates relative to wild-type *Mtb* were observed, strongly suggesting Mpa's involvement in the proteasomal delivery of pupylated substrates (Pierce *et al.*, 2008). This was further corroborated by *in vitro* studies of Mpa activity done by Striebel and coworkers, who only observed degradation of a GFP-pupylated protein, in the presence of Mpa (Striebel *et al.*, 2010). Therefore, pupylation of a target protein is specifically catalysed by PafA and degradation was dependent on the presence of Mpa ATPase - suggesting a highly regulated system.

1.3 Proteinopathy therapy

Mismanagement of proteins in cells can have deleterious effects, often leading to protein related diseases such as cancers and various neurodegenerative disorders e.g., Alzheimer's disease (Wang *et al.*, 2020). In such cases, therapeutic intervention is sought to help mitigate the resulting effects and restore cellular homeostasis. One such strategy is the use of small chemical molecules that are strategically designed to modulate target proteins (Gerry and Schreiber, 2018). These molecules often inhibit the target protein by binding to receptor active/allosteric sites, consequently disabling protein function. A vast majority of clinically used agents use protein inhibition as the underlying principle for therapy, however, this mode of action is faced with two major limitations i.e. (1) A high dosage is typically required for pharmacological relevant inhibition, which may induce undesired off-target effects and, (2) protein inhibition

via binding to active or allosteric site excludes much of the human proteome (~80%) which is classified as undruggable (Bondeson *et al.*, 2015).

Another therapeutic avenue that has garnered significant interest and success is the genetic and chemical knockdown strategies. Genetic knockdown of proteins involves the use of technologies such as antisense oligonucleotides, RNA interference (RNAi) and CRISPR/Cas9 gene editing (Walters *et al.*, 2016). Although the strategy seemingly overcomes the “druggability” limitation offered by chemical inhibitors, it faces many challenges such as possible off-target effects, challenges in delivery efficacy and poor metabolic stability. On the other hand, chemical knockdown strategies typically involve the use of a small bifunctional molecule which target and bind a protein of interest, whilst simultaneously recruiting the natural degradation system of the cell i.e., induced protein degradation (Bondeson and Crews, 2017). The high specificity associated with the chemical knockdown strategy has elicited significant attention, resulting in the development of specific non-genetic IAP-based protein erasers (SNIPERs), and chaperone-mediated protein degradation (CHAMPs), and proteolysis-targeting chimera (PROTACs) (Delpont and Hower, 2019).

1.3.1 Induced protein degradation in Eukaryotes

1.3.1.1 Specific non-genetic IAP-based protein erasers (SNIPERs)

The inhibitor of apoptosis (IAP) family of proteins is a class of proteins that help maintain cellular health by acting as apoptosis inhibitors. IAPs are commonly overexpressed in cancer cells and are often used to monitor cancer progression in patients. The cellular IAP1 (cIAP1), in particular, ubiquitinates the receptor interacting protein kinase 1 (RIPK1) via its E3 ligase-like RING finger domain, consequently inhibiting native RIPK1 programmed necrosis activity. Seen as a potential therapeutic target, studies targeting cIAP1 have been explored. One such strategy has been the use of drugs such as bestatin, a clinically relevant aminopeptidase inhibitor, that promote cIAP1 autoubiquitination hence resulting in its proteasomal removal. The induced removal of cIAP1 has been linked to antitumor activity which is showcased in phase III clinical trial studies by Ichinose and coworkers, where bestatin resulted in the prolonged survival of stage I squamous-cell lung carcinoma patients (Ichinose *et al.*, 2003). ME-BS, a bestatin homologue, and other bestatin analogues have also been recruited for the selective degradation of cIAP1. A study by Sekine and colleagues demonstrated significant

selective down-regulation of cIAP1 by ME-BS, bestatin, and bestatin esterified analogs via a ubiquitin-proteasome dependant pathway in the HT1080 fibrosarcoma cell-line (Sekine *et al.*, 2008). Five of the bestatin analogs BE01, BE14, BE15, BE32, and BE33, observed significant cIAP1 degradation and increased HT1080 fibrosarcoma cell-line anti-Fas (CH11) antibody dependent apoptosis. Additionally, the group observed comparable cIAP1 induced degradation by ME-BS at a 10 x lower concentration than bestatin. Another study by Bertrand and coworkers demonstrated potent antiproliferative activity of cIAP1, cIAP2, and XIAP-binding compound AEG40730 on two cancer cell lines i.e., MDA-MB-231 and SKOV3, displaying an $IC_{50} < 5$ nM (IC_{50} , drug concentration for half-maximal effect) (Bertrand *et al.*, 2008).

Small molecule compounds that bind cIAP1 have been recruited in the development of heterobifunctional molecules i.e., SNIPERs, that hijack the E3 ligase-like ubiquitination activity of cIAP1 for targeted proteasomal degradation. Unlike other E3 ligase recruiting drugs, some SNIPERs additionally facilitate cIAP1 degradation through induced autoubiquitination hence presents a synergistic approach towards antitumor activity. This is demonstrated in a study by Naito and coworkers who developed a series of highly specific bestatin-based SNIPERs by tethering the cIAP1 binder to a protein target ligand.

1.3.1.2 PROTACs

PROTACs, which have received intense scientific interest over the last two decades, are small heterobifunctional molecules made up of two efficient, distinct binding moieties connected by a chemical linker. These molecules are specially designed to target specific proteins for 26S proteasomal degradation using the UP-cascade system. Of the binding moieties, one is a protein binding ligand, aptly termed the “warhead”, and the other an E3 ligase-binding ligand (the recruiter) (Liu *et al.*, 2020). PROTACs, like the other degraders are classified as proximity-inducing drugs and function by binding both the E3 ubiquitin ligase and the target protein, essentially catalysing polyubiquitination of the target protein for eventual 26S proteasomal degradation (Figure 1.3). Upon its inception, PROTAC technology opened a whole new avenue in drug development against proteinopathies. This, of course, is due to several advantages displayed by PROTAC molecules compared to most modulating drugs. The most attractive of these, is the apparent ability of PROTACs to target the large number of proteins previously characterized as undruggable by typical modulatory agents. A classic example of this is the established oncogene c-Myc, which despite its attractiveness as a cancer therapy target

(expressed in >70% of cells), had been characterized as undruggable. This is due to the intrinsically disordered nature of c-Myc, expressing a structure that lacks the typical target sites for conventional modulatory effect (Madden *et al.*, 2021). Another major advantage of PROTAC technology is the complete removal of a target protein in contrast to modulatory agents that are required to constantly occupy their target modulatory sites for activity (Schneekloth *et al.*, 2008). This overcomes the common resistance issues often experienced by modulatory drugs such as inhibitors, lowering the likelihood of resistance by removing proteins altogether. The catalytic nature of a PROTACs enables their use at low concentrations where a single PROTAC molecule can effectively facilitate the degradation of several target proteins.

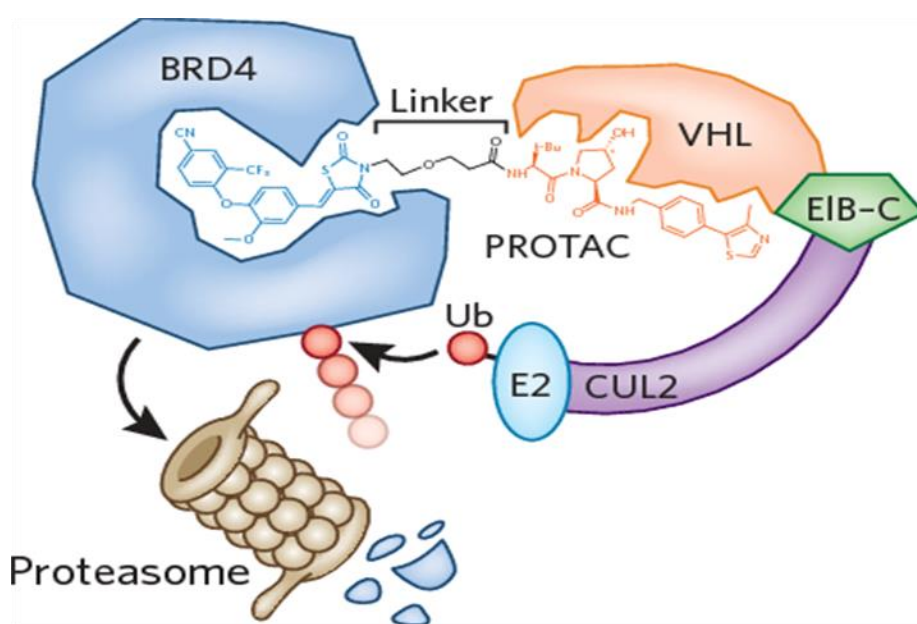


Figure 1.3. PROTAC mediated protein degradation. Here, the PROTAC links the target protein BRD4 with the VHL E3 ubiquitin ligase forming a ternary complex (BRD4:PROTAC:VHL). This facilitates polyubiquitination resulting in BRD4 being tagged for proteasomal degradation.

E3 ligases form critical elements of PROTAC induced proteolysis, recruited to facilitate polyubiquitination of the bound target protein for subsequent 26S proteasomal degradation. Estimates indicate that the human proteome expresses more than 600 types of E3 ligases that maintain the highly substrate-specific UP system (An and Fu, 2018). The cereblon (CRBN)-CRL4 (Cullin-RING) complex and the Von Hippel Lindau (VHL)-CRL2 complex are commonly recruited E3 ligases for targeted protein degradation.

1.3.1.2.1 Cereblon-engaging PROTACs

Cereblon is a protein ubiquitously expressed in the cytoplasm and nucleus of cells, functioning as the protein-binding motif of the CRL4 E3 ligase complex which consists of damaged DNA-binding protein 1 (DDB1), cullin-4A (CUL4A), and regulator of cullins 1 (ROC1)) (Wada *et al.*, 2016). Drug modulation of cereblon began with the invention and clinical use of the drug thalidomide, which was later discovered to have teratogenic activity and quickly withdrawn from the market (Fischer *et al.*, 2014). Subsequently, phthalimide-derived drugs thalidomide, lenalidomide, and pomalidomide later emerged as potent IMiDs against multiple myeloma and were thus repurposed for their anticancer activity (Rehman *et al.*, 2011). These drugs were reported to decrease CRL4^{CRBN} dependent ubiquitination of native substrates whilst simultaneously promoting ubiquitination of new protein substrates such as casein kinase 1 α and the transcription factors IKZF1 and IKZF3 for subsequent degradation (Krönke *et al.*, 2014; Petzold *et al.*, 2016). IMiDs have also been exploited for the development of cereblon-recruiting PROTACs, serving as the E3 ligase recruiting ligand (Ito and Handa, 2020).

Recently, Mu and colleagues developed and demonstrated a thalidomide derived PROTAC HBL-4, which displayed potent dual activity against the well-established therapeutic targets bromodomain 4 (BRD4) and polo-like kinase 1 (PLK1) in acute myeloid leukemia (Mu *et al.*, 2020). HBL-4 displayed high potency for both BRD4 and PLK1 proteins in a MV4-11 cell line observing a DC₅₀ (drug concentration to achieve 50% degradation of the target protein) of 5 nM and 15 nM, respectively. At 40 nM, HBL-4 was reported displaying near complete degradation of both proteins in two other cell lines i.e., MOLM-13 and KG1, within 24 h. Additionally, HBL-4 was shown to significantly suppress the oncogene c-Myc in a MV4-11 cell line, suggesting a synergistic approach towards antiproliferative activity. Similarly, Ling and co-workers observed significant BRD4 degradation and c-Myc downregulation in primary human thyroid carcinoma cells, using the cereblon recruiting PROTAC ARV-825 at nanomolar concentrations (He *et al.*, 2020).

Another BRD4 targeting PROTAC, dBET1, was designed by Winter and collaborators by linking the BRD4 and cereblon ligands JQ1 and thalidomide, respectively (Winter *et al.*, 2015). dBET1 displayed antiproliferative activity against a MV4-11 leukemia cell line *in vitro* and in a mouse model. In 2017, novel cereblon recruiting PROTACs derived from BET targeting γ -

carboline compounds, were synthesised and studied by Zhou and colleagues on their activity against BRD4 in RS4;11 cell line. Two of the PROTACs, compound 21 and 23, displayed an impressive degradation of BRD4 at a sub-nanomolar DC₅₀ concentration of < 0.3 nM in just 3 h.

The CRL4^{CRBN} E3 ligase complex has proven to be an efficient target for PROTAC mediated protein degradation. Displaying one of the most potent degradation activities currently recorded (at DC₅₀ <0.3 nM). Cereblon-recruiting PROTACs elicit interest in their application to degrade disease-related protein targets. Importantly, incorporation of IMiD molecules (pthalimide derivatives, etc) or other inhibitor compounds in PROTAC development needs careful consideration due to the high possibility of the developed drug retaining inhibitory activity, an undesired feature that would result in off-target effects (Delport and Hewer, 2019). This is showcased in a study by Schiedel and coworkers, who developed a CRL4^{CRBN} E3 ligase complex recruiting PROTAC by tethering thalidomide to a potent inhibitor of a cancer therapeutic target sirtuin2 (Sirt2), SirReal3. *In vivo* analysis of the PROTAC showed potent inhibitory activity towards Sirt2 with an IC₅₀ value of 0.25 µM.

1.3.1.2.2 VHL-engaging PROTACs

The VHL containing E3 ligase complex (CRL2^{VHL}) is another E3 ligase commonly recruited for small molecule mediated protein degradation (Ishida and Ciulli, 2021). Conceptualised in Figure 1.3., the VHL protein functions as the substrate recognition unit of the E3 ligase, recognising and binding relevant proteins for recycling via the UP cascade.

Initially, CRL2^{VHL} recruiting PROTACs were originally designed as bifunctional peptides that bind the VHL protein using the cell permeable peptide ALAPYIP, resulting in the first generation PROTAC drugs that facilitate targeted protein degradation without the need for microinjection (Pettersson and Crews, 2019). The discovery of the VHL peptidic ligand (2S,4R)-1-[(2S)-2-amino-3,3-dimethylbutanoyl]-4-hydroxy-N-[(1S)-1-[4-methyl-1,3-thiazol-5-yl]phenyl]ethyl]pyrrolidine-2-carboxamide;hydrochloride (AHPC), is considered the most important point in the development of PROTAC technology, paving the way for the progress made to date. The CRL2^{VHL} recruiting PROTAC library has since expanded, incorporating small molecule chemical compounds that display inducible target degradative activity using the

CRL2^{VHL} quality control system. Drug discovery of VHL-recruiting PROTACs has reported activity against a range of established oncogenic proteins particularly BRD4. For instance, Ciulli and coworkers reported BRD4 degradation using MZ1-3, a PROTAC designed by tethering the pan-BET selective bromodomain inhibitor JQ1 to a CRL2^{VHL} E3 ligase ligand, AHPC. Complete degradation of BRD4 was observed for all three compounds at 1 μ M, suggesting a < 1 μ M DC₅₀ concentration (Ciulli *et al.*, 2015). Further work by Ocaña and colleagues displayed MZ1 downregulation of BRD4 expression in JQ1-resistant triple negative cancer cell lines. Additionally, *in vivo* studies demonstrated MZ1 anti-tumor activity in a JQ1-resistant xenograft model via BRD4 degradation, displaying MZ1 potential in clinical application (Ocaña *et al.*, 2019).

Recently, a study was done by Carreira and coworkers which saw the development of a bi-stable persistent CRL2^{VHL} recruiting photoPROTAC derived from the BRD2/3/4 degrader ARV-771, (Carreira *et al.*, 2019). By using an ortho-F4-azobenzene linker between the PROTAC's ligands, the group were able to either activate or deactivate the PROTACs activity using 415 nm and 530 nm LED irradiation, respectively. Interestingly, the active form "trans-photoPROTAC" displayed selective degradative activity for BRD2 in a Burkitt lymphoma Ramos cell line, contrary to its parent ARV-771. This on/off switch technology for selective degradation activity displayed appeal for therapeutic application in protein related diseases. Dragovich and coworkers developed the first CRL2^{VHL} recruiting PROTAC, GNE-987, displaying an impressive 30 pM DC₅₀ concentration for BRD4 in EOL-1 cells and significant inhibition of c-Myc expression in the MV-4-11 cell line (Dragovich *et al.*, 2020).

Despite the high potential therapeutic application of PROTACs, the technology faces several challenges. These, for the most part, generally stem from the typical high molecular weight of PROTAC molecules often resulting in poor permeability and bioavailability (Cecchini *et al.*, 2021). Typically lacking drug-likeness, pharmacokinetics ranks as the greatest obstacle in the success of PROTAC technology in cancer therapy. This is showcased by the GNE-987 PROTAC, which significantly degraded BRD4 at picomolar concentrations but displayed poor *in vitro* metabolic stability and *in vivo* pharmacokinetics (Dragovich *et al.*, 2020). The team managed to overcome this shortcoming by conjugating an antibody, to the GNE-987 forming a GNE-987-antibody conjugate which improved both the stability and pharmacokinetics (Dragovich *et al.*, 2020). Various xenograft mouse model experiments using the GNE-987-

antibody conjugate displayed dose and antigen dependent anti-tumor activity (Dragovich *et al.*, 2020).

Excitingly, the increased focus in the development of new PROTACs especially by larger pharmaceutical agencies like Arvinas and Pfizer has yielded in the production of quality PROTAC drugs. Arvinas, in particular, now demonstrates the development of stable, bioavailable PROTAC drugs that additionally showcase favourable pharmacokinetics. This has resulted in an increased understanding in drug development, providing new “Arvinas rules” to guide future PROTAC development. Two PROTAC molecules designed by Arvinas therapeutics have recently entered human clinical trials with more expected to follow as the trials progress. Two of the drugs, ARV-110 and ARV-471 (Figure 1.4), have shown encouraging results so far in phase I/II clinical trials against prostate and breast cancer, both displaying favourable relevant anti-tumor activity (Mullard, 2019; Li and Song, 2020; Qi *et al.*, 2021). Currently, the two PROTACs are anticipated to enter phase III trials against metastatic breast cancer. The success of PROTACs so far, presents an enticing idea of developing similar drugs that target other proteasomal degradation cascades.

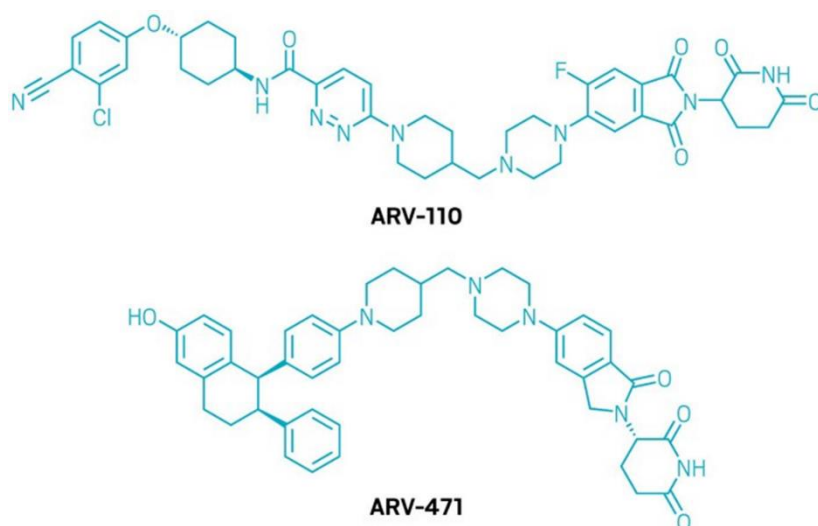


Figure 1.4. Phase I/II Arvinas clinical PROTACs

1.3.2 Induced protein degradation in prokaryotes

Modulatory agents that inhibit essential bacterial proteins for apoptosis are usually employed for therapeutic effects but are often challenged by the rapid development of resistance. This is due to the combined effects of low generation time and high mutation rates typically associated with bacterial populations, resulting in the increased likelihood of drug resistance development (Wilson *et al.*, 2016). Moreover, some bacterial species possess the ability to confer drug resistance to other bacterial species via horizontal gene transfer of drug resistance genes, enabling the emergence of multidrug resistant bacteria that hurdle available treatment options and presents a major concern in public health (Sun *et al.*, 2019).

To help mitigate resistance issues associated with classic drug inhibitors like antibiotics, drug induced targeted protein degradation in bacteria for antiproliferative effects has been under discussion (Dong *et al.*, 2021; Gopal and Dick, 2020) and only recently been explored in a study by Morreale and coworkers. The team developed novel heterobifunctional PROTAC-like molecules i.e., BacPROTACs, that showed significant *in vitro* and *in vivo* protease selective degradation of a model protein substrate monomeric streptavidin (mSA) (Morreale *et al.*, 2021). By linking the ClpC:ClpP (ClpCP) protease substrate receptor ligand phosphorylated arginine residues (pArg) to the mSA-binding ligand biotin, the group developed BacPROTAC molecules that facilitated selective mSA degradation by the ClpCP protease. One of the developed chimeric molecules, BacPROTAC-1, resulted in near-complete degradation of mSA at 1 μ M in a *Bacillus subtilis* *in vitro* study. This was also observed in *Mycobacterium smegmatis* *in vitro* studies where a 1 μ M BacPROTAC-1 similarly induced near complete degradation of mSA by hijacking the ClpCP homolog ClpC1P1P2. The group additionally developed cyclic antibiotic cyclomarin A (CymA)-based BacPROTACs, displaying high specificity for the *Mycobacteria* ClpC1P1P2 protease reducing possible off-target effects. BacPROTAC-2, synthesized by tethering biotin to a peptidic CymA analog, resulted in the near complete removal of the model mSA protein. The team additionally reported BacPROTAC induced degradation of bromodomain 1 (BRD1), a transcription regulator protein that, when downregulated, sensitises cancer cells to apoptosis (Fryland *et al.*, 2016; Klein *et al.*, 2018; Winter *et al.*, 2015). This chimeric degrader i.e., BacPROTAC-3, was synthesised by joining the peptidic CymA analog used in BacPROTAC-2 design with the JQ1 ligand typically recruited in PROTAC design. BacPROTAC-3 *in vitro* studies using *Mycobacterium smegmatis*

resulted in significant ClpC1P1P2 mediated BRD1 degradation with 1 μ M resulting in almost complete BRD1 removal.

Throughout the study, Morreale and colleagues not only demonstrated BacPROTAC selective degradation of protein targets, but also showed that degradation is dependent on the formation of a ternary complex for activity like in eukaryotic PROTACs. This observed induced degradation of target proteins in a prokaryotic environment i.e., *Mycobacteria*, elicits interest in the exploration of other endogenous degradation systems of the bacterium for targeted degradation. Pupylation, the post-translational protein modification system limited to *Mycobacteria* and other Actinobacteria (Pierce *et al.*, 2008), presents an attractive and innovative target for PROTAC-like drug development. Akin to ubiquitination, target proteins are flagged for Mpa-20S proteasomal degradation using the small protein pup, ligated to target proteins using the E3 ligase functional analogous enzyme pup-ligase (proteasome accessory factor A; PafA) (Figure 1.2).

1.3.3 Pupylation and tuberculosis

The pupylation cascade is found in one of the deadliest organisms, *Mtb*, the agent of the chronic disease tuberculosis (TB). TB is the leading cause of death by single infection in South Africa and counts amongst the top ten diseases with the highest mortalities worldwide (Loveday *et al.*, 2019). The traditional treatment option typically involves a combination of antibiotic drugs for a fixed period, often resulting in a range of side effects including hepatitis and dyspepsia (Janssen *et al.*, 2019). Over the years, mismanagement of anti-*Mtb* antibiotic drugs has contributed to the emergence of multi-drug resistant *Mtb* and recently, extensively drug-resistant and completely drug-resistant *Mtb* strains (Jiang *et al.*, 2018). Pupylation has been reported to contribute to *Mtb* persistence inside host cells, facilitating resistance against the various chemical stresses of macrophages (Özcelik *et al.*, 2012). Through extensive research, resistant mechanisms towards currently deployed drugs have been characterized, most reports observing reduced drug affinity on target binding sites due to single point mutations (Blanchard, 1996.; Palomino and Martin, 2014.; Ramaswamy and Musser, 1998.; Telenti *et al.*, 1993). Rifampicin resistance (a common anti-*Mtb* first-line drug), for instance, has been linked to alterations of the drugs target i.e., the β -subunit of the DNA-dependent RNA polymerase (RNAP) due to single point mutations on the associated *rpoB* gene (Figure 1.5). Discoveries by Telenti and coworkers observed that substitution(s) in highly conserved amino acids of the

RNAP β -subunit, encoded by the RNAP β -subunit gene *rpoB*, are responsible for a ‘single step’ high level resistance in *Mtb* (see table 1.1 for other gene mutations associated with resistance of other anti-*Mtb* drugs) (Telenti et al, 1993).

Table.1.1 Summary of *Mtb* mutated genes conferring resistant against first-line defence anti-*Mtb* drugs

| Drug | Associated mutated gene or mutation |
|--------------|---|
| Rifampin | <i>rpoB</i> |
| Isoniazid | <i>katC</i> , <i>inhA</i> , <i>oxyR</i> , <i>ahpC</i> , <i>furA</i> |
| Pyrazinamide | <i>pncA</i> , IS6110 insertion |
| Ethambutol | <i>embB</i> |
| Streptomycin | <i>rrs</i> , <i>rpsL</i> |

(Gillespie, 2002)

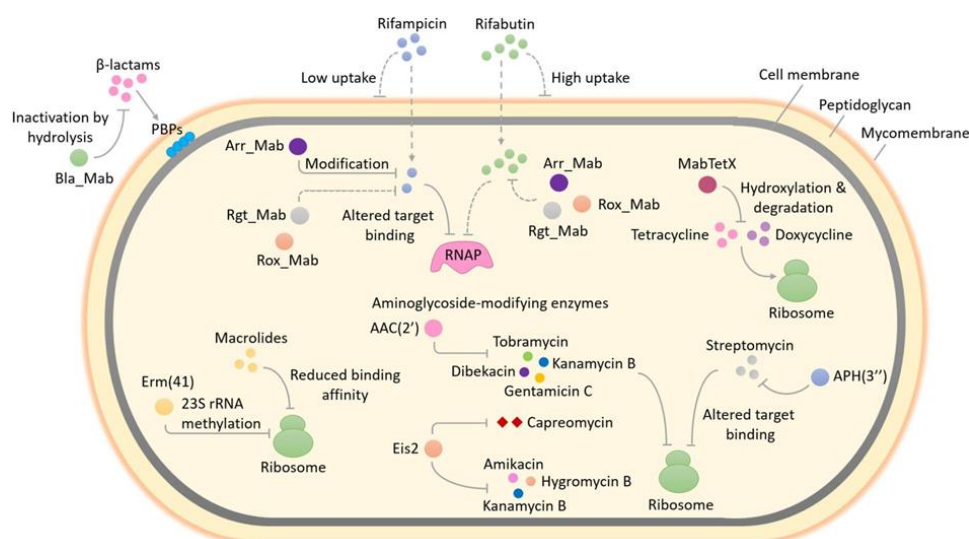


Figure 1.5. *Mycobacterium* drug resistance mechanisms. Various resistance mechanisms against modulatory drugs are shown including rifampicin resistance via the modification of its target site RNAP (Luthra et al., 2018).

1.4. Study rationale

One such strategy, is the development of drugs that possess the innovative ability to recruit the pupylation cascade of *Mtb*, much like PROTAC technology in the UP system of eukaryotes. The development of this novel type of BacPROTAC, would confer the ability to abrogate *Mtb* proliferation by targeting the pathogens essential proteins. The RNAP of *Mtb* presents as an initial target for the development of these prokaryotic degraders. This essential protein, an already established anti-*Mtb* target, has a well characterised binding ligand rifampicin (i.e., the warhead). To complete the warhead:linker:pupylation recruiter bifunctional molecule ligands that are capable of recruiting PafA, the sole ligase mediating pupylation, are therefore needed. Identification of PafA-binding ligands for the ultimate development of novel degraders would open a new avenue of TB therapy, providing insight on if induced protein degradation could be recruited in an anti-TB effort. Additionally, future recruitment of anti-*Mtb* antibiotics as the protein targeting motif of the degraders would show if these antibiotics could be repurposed and the developed degrader effective against resistant *Mtb* strains.

1.5 Aims and Objectives

The proposed study aimed to identify novel ligands that bind PafA. This was achieved through the fulfilment of the following objectives:

1. PafA was recombinantly expressed, purified, and characterised
2. A non-specific inhibitor screening assay was designed and optimised
3. Compounds from a small-molecule chemical library were screened for non-specific inhibiting activity
4. A PafA thermal shift assay (TSA) was designed and optimised
5. The small-molecule compound library was screened for potential PafA-binding ligands using the TSA

CHAPTER 2: MATERIALS AND METHODS

To identify PafA-binding ligands, PafA was recombinantly expressed in *E. coli* cells, solubilised in the presence of SDS detergent, purified via immobilised metal affinity chromatography (IMAC), and refolded using dialysis. Refolded PafA was thereafter screened for binders against a small-molecule chemical library using an optimised PafA TSA.

2.1 Recombinant expression of PafA

Vector constructs housing genes for the pupylation proteins PafA and pup, pET24b(+)-PafA and pET24b(+)-PupEPafA, respectively housed in DH5 α (EHD224) and DH5 α (EHD825) *E. coli* cell lines, were generously provided on filter disks by Professor Heran Darwin from the Department of Microbiology, New York University, USA (Cerdeira-Maira *et al.*, 2020). Due to shipment issues, the recombinant cells could not be revived on culture medium.

2.1.1 Vector recovery

To recover the recombinant pET24b(+)-PafA and pET24b(+)-PupEPafA vectors from the dead recombinant *E. coli* cells, filter disks were aseptically transferred onto sterile nutrient agar plates, 60 μ L Tris-EDTA (TE) buffer (10 mM Tris-HCl; 1 mM EDTA, pH 7.4) added on top of the filter disks to resuspend vector DNA, and incubated at room temperature for 10 min. Following incubation, vector DNA was recovered into sterile microcentrifuge tubes via pipette as TE buffer-DNA solutions, and subsequently analysed at 260 nm and 280 nm using the Nanodrop 2000 (Thermo Fisher Scientific, Massachusetts, USA).

2.1.2 Competent cell preparation and transformation

Non-recombinant JM109 (DE3) and B121 (DE3) *E. coli* cells were prepared for plasmid transformation via the CaCl₂-heat shock transformation protocol (Sambrook *et al.*, 2001). Initially, non-recombinant *E. coli* cells from a previously prepared 50% (v/v) glycerol stock solution were streaked onto fresh 2xYT agar plates (1.6% (w/v) tryptone, 1.0% (w/v) yeast extract, 0.5% (w/v) NaCl, and 1.5% (w/v) bacteriological agar), and cells grown overnight at 37 °C. A fresh isolate colony from the overnight growth plates was selected and subsequently transferred aseptically into 10 mL 2xYT broth medium and grown overnight at 37 °C and 200 rpm. The overnight culture was diluted 1:100 with 2xYT broth medium and grown to an

OD_{600 nm} of 0.3 - 0.4 at 37 °C. The resulting culture was transferred into ice cold sterile centrifuge tubes, incubated on ice for 10 min, and cells recovered by centrifugation (5000 x g, 10 min, 4 °C). The supernatant was discarded, and the cell pellet resuspended in 40 mL sterile CaCl₂ (60 mM CaCl₂, 10 mM HEPES, pH 7.0). The solution was centrifuged (5000 x g, 10 min, 4 °C), and the resulting pellet resuspended in 2 mL of the CaCl₂ solution, yielding competent cells.

To transform the cells, 2 µL of the TE buffer-impure DNA solution (~ 1129 ng/µL DNA; A_{260nm} / A_{280nm} = 1.70, A_{260nm} / A_{230nm} = 0.68) was added to 20 µL of the CaCl₂ competent cells and the solution incubated on ice for 30 min, followed by heat shocking at 42 °C for 90 s and further ice incubation for 2 min. For increased transformation efficiency, the cells were incubated for 1 h in 80 µL pre-warmed super optimal cataboliser (SOC) medium [2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 10 mM MgCl₂, 20 mM glucose, 2.5 mM KCl, and 10 mM MgSO₄] at 37 °C with gentle shaking. Thereafter, the resulting cell culture was plated onto 2xYT agar plate(s) (50 µg/mL Kanamycin supplement (Life Technologies, California, USA)) and cells incubated overnight at 37 °C

Of the resulting overnight colonies, four were selected to assess transformation efficacy. Each colony was inoculated into 10 mL 2xYT broth culture (100 µg/mL Kanamycin) and grown overnight at 37 °C and 200 rpm. A 750 µL 25% (v/v) glycerol stock solution was made for each overnight culture, leaving sufficient culture for plasmid isolation using the QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany), following the manufactures protocol. The isolated vector, pET24b(+)-PafA and pET24b(+)-PupEPafA, were analysed via single restriction digestion to assess plasmid size (6.6 kbp and 6.7 kbp expected sizes, respectively) using 0.5 U *Nde*I, 1.25 X CutSmart buffer (BioLabs, Massachusetts, USA), and 1 µg DNA in a 40 µL final reaction volume. Restriction digestion products were analysed on a 1% (w/v) agarose gel run at 80 V and stained with 0.005% (v/v) ethidium bromide (Promega, Wisconsin, USA) using an O'GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific Massachusetts, USA).

2.1.3 Polymerase Chain Reaction (PCR) and Sequencing of PafA and PupE

Appropriate primers were needed to amplify PafA and the PupE genes from the respective pET24b(+)-PafA and pET24b(+)-PupEPafA plasmid constructs. The universal T7 promotor and T7 terminator primer set (see sequence in Table.2.1) were selected to amplify PafA in the

pET24b(+)-PafA construct. For PupE amplification in the pET24b(+)-PupEPafA dual expression system, primers were synthesised that specifically target the PupE gene cloned into the BglII and XbaI sites of the construct. The nucleotide sequence encoding the pup protein tag was sourced from the NCBI database (GeneID: 888788) and was used to design appropriate primers (Table 2.1) using SnapGene 5.2.3 for use in both PCR and Sanger sequencing. Primer synthesis and subsequent Sanger sequencing of both plasmid constructs using the relevant primers was done by Inqaba Biotec (Inqaba Biotec, Pretoria, South Africa).

Table 2.1 Recombinant plasmid PCR primers.

| Plasmid | Sequences (5'- 3') | |
|------------------------|--|--|
| pET24b(+) PafA | Forward primer (T7 promotor) (T _m = 50 °C) | Reverse primer (T7 terminator) (T _m = 57 °C) |
| | TAATACGACTCACTATAGGG | GCTAGTTATTGCTCAGCGG |
| pET24b(+)-PupE PafA | Forward primer (synthesised) (T _m = 63 °C) | Reverse primer (synthesised) (T _m = 62 °C) |
| | ACTGAATTCATGGCGCAAGA GCAGACCAAG | ACTGGATCCTCACTGTCCG CCCTTTTGGAC |

*T_m, primer melting temperature

Both PafA and PupE were amplified from their respective plasmid constructs using 1 x *Thermus aquaticus* Master Mix (BioLabs, Massachusetts, USA, Catalogue No: M0270L), 0.2 µM forward and reverse primers, 10 ng vector constructs, and PCR reactions made up to 50 µL final assay volume with MilliQ water. All PCR reactions were performed in the G-Storm GS1 Thermocycler (G-Storm Ltd, Somerset, England) using the cycle conditions described in Figure 2.1, and the PCR products analysed on a 1% (w/v) agarose gel electrophoresis gel stained with 0.005% (v/v) ethidium bromide using the GeneRuler 100 bp and O'GeneRuler 1 kb DNA ladders.

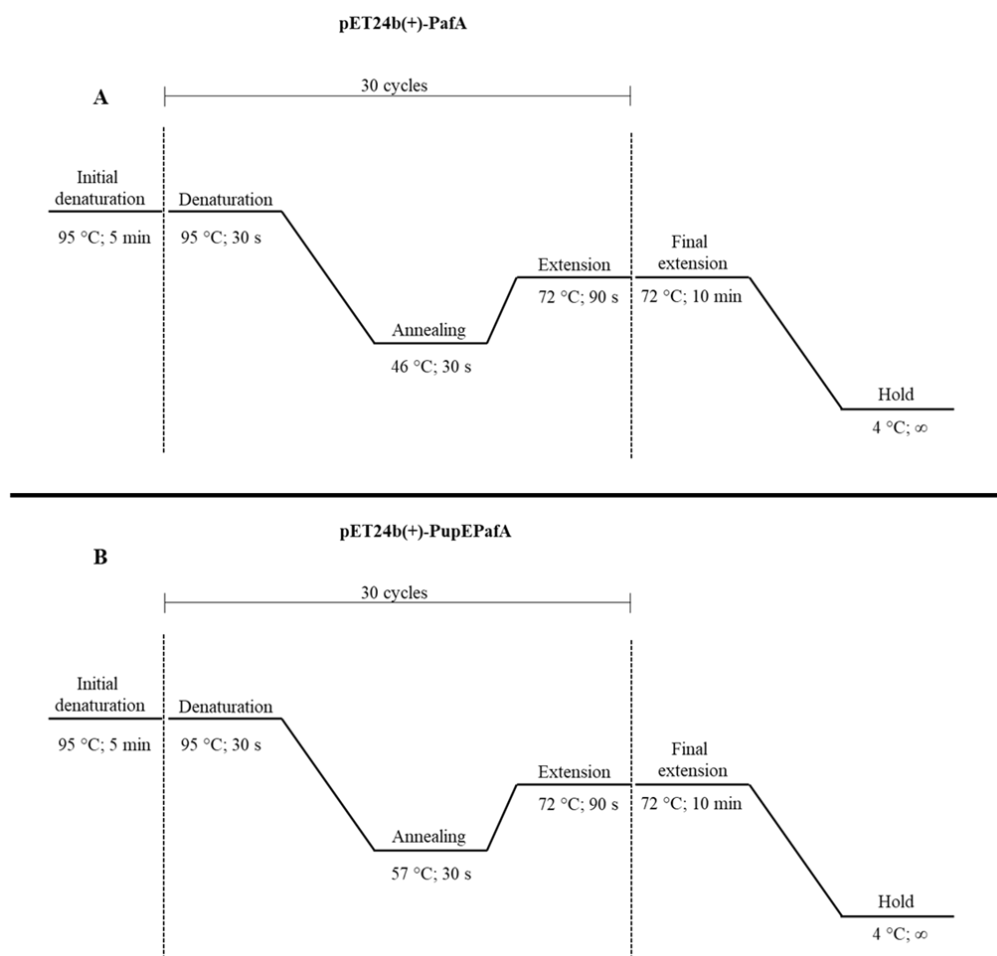


Figure 2.1. PCR reaction conditions for the amplification of (A) PafA and (B) Pup from the pET24b(+)-PafA and Pet24b(+)-PupEPafA constructs, respectively.

Sanger sequencing data received from Inqaba Biotech was edited and aligned using Chromas 2.6.6 (Technelysium, Brisbane, Australia), MEGA11 (Pennsylvania State University, Pennsylvania, USA), and BioEdit 7.2.5 (BioEdit Company, California, USA). Aligned sequences were used to generate a consensus sequence, which was then translated using the online bioinformatics resource portal ExPASy (SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland). Each translated product was thereafter analysed on BLASTp via a two-sequence alignment against the respective *Mtb* protein sequence i.e., either PafA or Pup sourced from the NCBI database, and the alignment scores and E-values recorded.

2.1.4 Recombinant PafA expression

Recombinant pET24b(+)-PafA JM109 (DE3) and Bl21 (DE3) *E. coli* cells, and recombinant pET24b(+)-PupEPafA JM109 (DE3) *E. coli* cells were used for the expression of the 51-kDa PafA conjugated with a 6xHis-tag at the carboxy-terminal, as amended from Cerda-Maira *et al*, (2020). Initially, the recombinant cells were plated onto 2xYT (100 µg/mL Kanamycin) agar medium for 24 h at 37 °C, thereafter a single colony was transferred into 10 mL 2xYT (100 µg/mL Kanamycin) broth medium and incubated overnight (37 °C, 200 rpm). The overnight culture was then diluted 1:100 in fresh 2xYT (100 µg/mL Kanamycin) broth medium (including negative control samples) and the resulting culture incubated under previous conditions until reaching an OD_{600 nm} of 0.5 - 0.7. PafA expression was then induced using a 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) (Melford Laboratories, Ipswich, UK) concentration and cultures incubated for 4 h at 37 °C, 200 rpm. The negative control samples for each cell line included non-recombinant and recombinant *E. coli* cell cultures, grown in the absence of the inducer IPTG, and a non-recombinant *E. coli* cell culture exposed to 1 mM IPTG.

For whole protein profile analysis, 2 mL culture samples were centrifuged (5000 x g, 1 min, room temperature), and either stored at -20 °C or resuspended in a 600 µL 1:1 solution of phosphate buffered saline [PBS; 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, pH 7.4] (pH 7.4) and reducing treatment buffer (0.125M Tris-HCl buffer (pH 6.8); 4% (w/v) SDS; 20% (v/v) glycerol; 10% (v/v) 2-mercaptoethanol). Prior to reducing sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) analysis, all sample solutions were sonicated 15 x on ice (50% power; 30-s sonication, 30-s rest) using the Fisherbrand 120 sonic dismembrator (Thermo Fisher Scientific Massachusetts, USA) and boiled for 90 s. Samples were kept on ice until ready to load onto the gel.

2.1.5 SDS-PAGE and immunoblot analysis

Protein samples were analysed using 10% reducing SDS-PAGE gels and Spectra Multicolour Broad Range Protein Ladder (Thermo Fisher Scientific, Massachusetts, USA), run at 18 mA/gel and stained with Bromophenol blue dye (Thermo Fisher Scientific, Massachusetts, USA). For immunoblot analysis, proteins from an unstained 10% SDS-PAGE gel were transferred onto a nitrocellulose membrane (Pall Corporation, New York, USA) overnight (10 V, room

temperature), and PafA detection done following a protocol by Towbin *et al*, (1979) with slight modifications. Briefly, the resulting nitrocellulose membrane was blocked in 5% (w/v) non-fat milk powder dissolved in a Tween-20 Tris buffered saline (TBSt; 20 mM Tris, 200 mM NaCl, 0.001% (v/v) Tween-20, pH 7.4) solution for 1 h at room temperature, and subsequently washed 3 x 5 min with the TBSt solution. Thereafter, the nitrocellulose membrane was incubated with the primary (1°) mouse anti-6xHis antibody solution (Catalog No.: R93025, Thermo Fisher Scientific, Massachusetts, USA), diluted 1:5 000 using 0.5% (w/v) bovine serum albumin (BSA)-TBSt, for 2 h. Following incubation, the blot was washed 3 x 5 min with TBSt and incubated for 1 h with a horse radish peroxidase (HRP)-conjugated goat anti-mouse IgG secondary (2°) antibody diluted 1:10 000 (Catalog No.: 31430, Thermo Fisher Scientific Massachusetts, USA) in 0.5% (w/v) BSA-TBSt. The blot was further washed 3 x 5 min in TBSt, and detection of protein bands done by enhanced chemiluminescence (ECL) (Novex, Massachusetts, USA). All SDS-PAGE gels and western blots were visualised using the G:BOX Chemi XR5 imagery system and the GeneTools 1.8.0 (Syngene, Cambridge, UK).

To determine protein solubility, 2 mL cell culture samples were pelleted (5000 x g, 5 min), resuspended in 600 µL PBS buffer supplemented with 750 µg/mL lysozyme (Thermo Fisher Scientific, Massachusetts, USA), and the resulting solution incubated at 37 °C for 30 min with gentle mixing. The lysozyme lysed samples were thereafter sonicated 15 x on ice (50% power; 30-s sonication, 30-s rest), centrifuged (12 000 x g, 30 min, 4 °C) and the supernatant soluble sample isolated from the insoluble pellet sample. The separated fractions were then prepared for reducing SDS-PAGE analysis by diluting the soluble fraction 1:1 in reducing treatment buffer or resuspending the insoluble pellet fraction in an equivalent volume (to the soluble fraction) of 1:1 reducing treatment buffer-PBS solution.

2.1.6 PafA solubilisation and purification

Two expression conditions were varied to assess their effect on PafA expression as inclusion body aggregates, namely, IPTG concentration and temperature. Initially, expression was assessed at 37 °C using IPTG concentrations of 0.1 mM, 0.5 mM, and 1 mM, sampling 2 mL every hour for 4 h to be analysed for PafA solubility via SDS-PAGE. Next, low temperature expression at 20 °C was investigated using the same IPTG concentrations (0.1 mM, 0.5 mM, and 1 mM) for 20 h. Following unsuccessful PafA soluble expression, PafA aggregates were solubilised using 1% (w/v) SDS following a protocol amended from Schlager *et al*, 2012. First,

PafA inclusion bodies were isolated from the soluble protein fraction of pET24b(+)-PafA recombinant expression culture (1 mM IPTG; 20 °C, 20 h) by centrifugation following lysis with lysozyme and sonication as described previously (section 2.1.5). Thereafter, the insoluble inclusion body fraction was washed 2 x with 50 mL 0.5% (v/v) Tween-20 in 1 x equilibration buffer (1xEB; 50 mM Na₂HPO₄, 300 mM NaCl) buffer solution to remove contaminating proteins (lipid and membrane-associated proteins), and PafA solubilised by the addition of 1% (w/v) SDS 1xEB buffer and subsequent 15 x sonication on ice (50% power, 30-s sonication, 30-s rest). Thereafter, samples were centrifuged (12 000 x g, 30 min, 4 °C) and the soluble supernatant sample isolated from the insoluble pellet sample. Each fraction was then analysed by 10% reducing SDS-PAGE as done previously (section 2.1.5).

PafA purification was carried out using IMAC on TALON resin (Co²⁺) (Takara Bio, Shiga, Japan) from the solubilised protein sample. Referencing the manufacturer instructions, the soluble fraction was applied to the prepared resin (equilibrated 3 x 10 min with 20 x bed volume 1xEB) and incubated at 4 °C, gently rotating using the RotoBot Programmable Rotator (Benchmark Scientific, Sayreville, USA) for 1 h. The resin was pelleted at 700 x g and the unbound sample carefully removed, reserving a sample for SDS-PAGE and immunoblot analysis. The resin bed was subsequently washed with 5 mM imidazole in 1xEB buffer (20 x bed volume), gently rotated for 10 min, and the resin pelleted at 700 x g. The supernatant was carefully removed, reserving a sample for downstream analysis (wash 1), and the resin bed washed again 4 x following the same protocol. After washing, the bound 6xHis-tagged protein was eluted by adding 250 mM imidazole in 1xEB (1 x bed volume), 5 min vortexing, and resin pelleted at 700 x g for 5 min. The supernatant was carefully removed into sterile tubes, and elution repeated until all bound 6xHis-tagged protein was eluted as determined by absorbance at 280 nm using the Nanodrop 2000 ($A_{280} < 0.2$) (Thermo Fisher Scientific, Massachusetts, USA). All samples were analysed by reducing SDS-PAGE and immunoblot following the previous protocol detailed in *Section 2.1.5*.

Following SDS-PAGE and immunoblot analysis, purified elute samples were dialysed against 100 x volume of 1xEB overnight at 4 °C using SnakeSkin Dialysis Tubing (Thermo Fisher Scientific, Massachusetts, USA) with gentle shaking, and the samples stored in sterile tubes for downstream experiments.

2.2 The detection of aggregating pan assay interfering compounds (PAINs) via a colorimetric enzyme assay

Before screening a 48-compound set from the DiverSET CHEMBRIDGE compound library generously provided by Doctor Mosebi from the Department of Life and Consumer Sciences (University of South Africa, South Africa) for potential PafA-binders, these compounds were screened for possible non-specific inhibiting properties. Non-specific inhibitor compounds typically exert inhibiting activity against more than one enzyme, and as such enzyme assays have been developed to screen for potential promiscuous inhibitors. These assays typically screen for inhibiting properties of compounds against a random protein's activity assay where "hit compounds" are flagged as potential promiscuous inhibitors. To this effect, an inexpensive and enzymatic *in vitro* assay was developed and optimised using the enzyme substrate duo HRP and its chromogenic substrate 3,3',5,5'-Tetramethylbenzidine (TMB).

2.2.1 Enzyme assay

The colorimetric assay was carried out using HRP (Thermo Fisher Scientific, Massachusetts, USA) dissolved in PBS supplemented with 0.001% Tween-20, and 50 ng/mL TMB dissolved in a 0.02% H₂O₂ 1:1 dimethyl sulfoxide (DMSO) and 150 mM citrate phosphate buffer (pH 5) solution. All experiments were performed in clear, flat-bottomed immuno nonsterile 96-well plates (Thermo Fisher Scientific, Massachusetts, USA) at a final assay volume of 200 µL, and spectrophotometric analysis done at 652 nm using a SpectraMax[®] ABS microplate reader (Molecular Devices, California, USA). Assays were carried out by the sequential addition of assay buffer, HRP, and TMB to their effective concentrations, and the resulting solution incubated for 10 min in the dark followed by spectrophotometric analysis at 652 nm. Assay optimisation was carried out by varying HRP concentration (5 ng/mL - 100 ng/mL) against the standard 50 ng/mL TMB. All assays were conducted in triplicate and each replicate adjusted from a no HRP reference control i.e., all assay components except hrp enzyme to account for possible background interference.

2.2.2 Non-specific inhibitor assay

A non-specific inhibitor detecting assay was developed using the optimised HRP-TMB assay and known aggregating inhibitors. Compound stocks of quercetin (Koch-Light laboratories Ltd, Haverhill, England); EGCG (MilliporeSigma, Massachusetts, USA); riboflavin (MilliporeSigma, Massachusetts, USA); tannic acid (Merck, Darmstadt, Germany) and 8-

hydroxyquinoline (Riedel-de Haen, Seelze, Germany) were prepared in DMSO and used at a final concentration of 10 μ M. Additionally, two negative control compounds caffeine (BDH Chemicals Ltd, Poole, England) and warfarin (Fluka AG, Buchs, Switzerland) were similarly used at 10 μ M. Sodium azide (BDH Chemicals Ltd, Poole, England) was included as a positive control (a known irreversible HRP inhibitor; $K_i = 1.47$ mM; K_i , inhibitor constant i.e., compound concentration for half maximum inhibition (Ortiz de Montellano *et al.*, 1988)), used at a final concentration of 1.5 mM. Experiments were initiated by preincubating HRP and inhibitor compound in PBS for 5 min in the dark. TMB substrate was subsequently added to visualise the reaction followed by 652 nm spectrophotometric analysis after 10 min incubation in the dark. All assays were conducted in triplicate and each replicate adjusted from a no HRP reference control i.e., all assay components except HRP enzyme to account for background interference. Percentage inhibition of HRP was calculated using the equation:

$$\%HRP\ Inhibition = \left(\left(\frac{(A652\ nm\ HRP\ DMSO\ control\ rxn - A652\ nm\ inhibitor\ rxn)}{A652\ nm\ HRP\ DMSO\ control\ rxn} \right) \times 100 \right)$$

Prior to screening the DiverSET CHEMBRIDGE library compounds via the developed non-specific inhibitor assay, the library was virtually screened for aggregating activity based on their chemical structure using the ZINC15 chemical database (<https://zinc15.docking.org/patterns/home/>). Thereafter, the 48-compound set was screened using the non-specific aggregating inhibitor detection assay at 10 μ M compound concentrations, and compounds displaying significant inhibiting activity against the HRP-TMB enzymatic assay (> 20% inhibition) were eliminated from downstream experiments.

2.3 Screening for potential PafA-binding ligands

To determine potential PafA-binding ligands, TSA was employed to screen the DiverSET CHEMBRIDGE library against purified PafA. In the TSA, a protein sample is gradually heated in the presence of a hydrophobic dye where, upon protein unfolding, the dye binds to the exposed hydrophobic regions resulting in increased fluorescence. Using fluorescence data from the protein thermal profile, protein melting temperature (T_m ; minimum temperature where 50% of a protein is unfolded) is extrapolated. Binding of a small molecule to proteins typically results in a positive shift of the proteins T_m i.e., stabilises the protein molecule thus needing more thermal energy to exert the same protein unfolding effects observed for the protein without

ligand. This shift in T_m is the principle used in early drug discovery to determine potential protein binding ligands, benefiting from its ease and high throughput.

2.3.1 PafA TSA optimization

Initially, a control TSA was carried out using the Protein Thermal Shift Starter Kit (Thermo Fisher Scientific (Catalog No.: 4462263), Massachusetts, USA), and SYPRO Orange (Thermo Fisher Scientific (Catalog No.: S6650), Massachusetts, USA) fluorescence dye. Reactions were made by adding 5 μ L PBS into a sterile 0.2 mL strip PCR tube, followed by the addition of 100 ng/ μ L control protein and 1 mM ligand (Table 2.2). A final volume of 17.5 μ L was made up using MilliQ water and the resulting solution gently mixed, centrifuged (10 300 x g for 10-s), and incubated on ice for 30 min. After incubation, 2.5 μ L from a prepared SYPRO Orange dye stock made in MilliQ deionised water was added into the solution to a final 5 x dye concentration, gently mixed, centrifuged (10 300 x g for 10-s), and the solution kept on ice and in the dark until thermal melt analysis. All assays were performed in duplicates and the thermal melt reactions carried out according to the cycle conditions shown in Figure 2.2, from 25 °C to 99 °C, and analysis done at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths using the QuantStudio 5 Real-Time PCR System and QuantStudio Design and Analysis software 2.6.0 (Thermo Fisher Scientific, Massachusetts, USA).

Table 2.2. Reaction setup for the control protein TSA

| Reagent* | Stock concentration | Volume (μL) | Final concentration |
|------------------------|----------------------------|-----------------------------------|----------------------------|
| PBS | 1 x | 5 | 0.25 x |
| MilliQ deionised water | - | 8.5 | - |
| Control protein | 1000 ng/ μ L | 2 | 100 ng/ μ L |
| Control ligand | 10 mM | 2 | 1 mM |
| SYPRO Orange | 40 x | 2.5 | 5 x |

*Reactions were prepared chronologically as listed

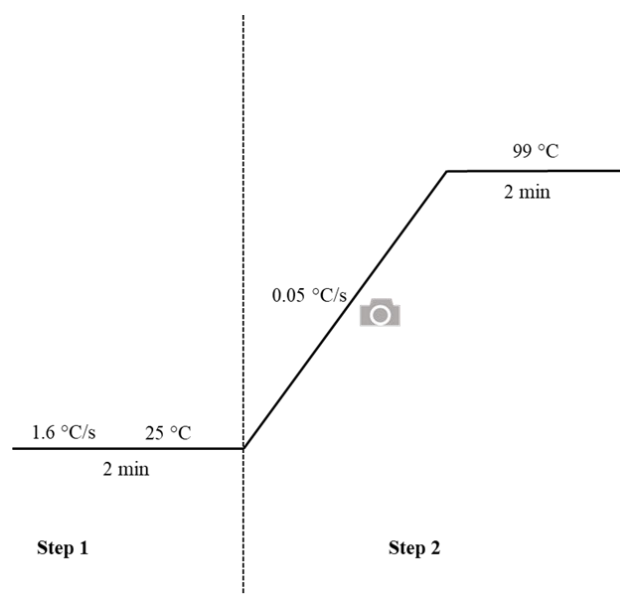


Figure 2.2. Thermal melt conditions for PafA thermal shift assays. Fluorescence analysis was done from 25 °C to 99 °C per 0.05 °C increment at 470 nm (\pm 15) excitation and 586 nm (\pm 10) emission wavelengths. The camera icon indicates the fluorescence analysis step of protein thermal melt reaction assays.

Before assessing PafA for potential binding ligands, the design and optimisation of a thermal shift assay (TSA) using the fluorescence dye SYPRO Orange (Thermo Fisher Scientific, Catalogue No: S6650), Massachusetts, USA) was also done. To this effect, a range of purified PafA concentrations (1 ng/ μ L, 2.5 ng/ μ L, 5 ng/ μ L, 7.5 ng/ μ L, 10 ng/ μ L, 50 ng/ μ L, 100 ng/ μ L, 150 ng/ μ L, 200 ng/ μ L, 250 ng/ μ L, and 300 ng/ μ L) were assayed against 1 x, 5 x, 10 x, 15 x, and 20 x, SYPRO Orange concentrations. Duplicate reactions were made up by adding 5 μ L PBS into a sterile 0.2 mL strip PCR tube, followed by the addition of an appropriate amount of purified PafA. A final volume of 17.5 μ L was made up using MilliQ water and the resulting solution gently mixed, centrifuged (10 300 x g for 10-s), and incubated on ice for 30 min. After incubation, 2.5 μ L from an appropriately prepared SYPRO Orange stock made in MilliQ deionised water was added into the solution, gently mixed, centrifuged (10 300 x g for 10-s), and the thermal melt reactions carried out using the cycle conditions shown in Figure 2.2, from 25 °C to 99 °C, and thermal melt analysis affected at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths.

PafA T_m was determined via the GrapPad Prism analysis software 9.2.0. Raw fluorescent data from the thermal melt reactions was initially exported as an Excel file from the QuantStudio Design and Analysis software 2.6.0, and thereafter fed into GraphPad Prism 9.2.0 for analysis using the non-linear regression: bell shaped, X is concentration parameters. PafA T_m was taken

as the mean $EC_{50} \pm 1$ standard deviation extrapolated from the generated thermal melt curves displaying goodness of fit R^2 -values ≥ 0.98 .

2.3.2 Screening for PafA-binders

To screen for possible PafA-binders from the DiverSET CHEMBRIDGE library, 10 μ M of each compound from the library was used to assess binding efficiency to PafA using the optimised PafA-SYPRO Orange thermal melt conditions i.e., 10 ng/ μ L PafA and 1 x SYPRO Orange. Reactions were made by the sequential addition of 5 μ L PBS, 8.3 μ L MilliQ deionised water, 4 μ L PafA (50 ng/ μ L 1xEB dialysed stock), and 0.2 μ L DiverSET CHEMBRIDGE compound (1 mM DMSO stock) in 0.2 mL strip PCR tubes, and the resulting solution gently mixed, centrifuged (10 300 x g for 10-s), and incubated for 30 min on ice (Table 2.3). After incubation, 2.5 μ L SYPRO Orange dye (8 x stock made in MilliQ deionised water) was added and the solution gently mixed, centrifuged (10 300 x g for 10-s), and kept in the dark until the thermal melt reactions were carried out using the cycle conditions shown in Figure 2.2.

Table.2.3. TSA Reaction setup for PafA-binding compounds

| Reagent* | Stock concentration | Volume (μL) | Final concentration |
|------------------------|--------------------------------|---------------------------------------|--------------------------------|
| PBS | 1 x | 5 | 0.25 x |
| MilliQ deionised water | - | 8.3 | - |
| PafA | 50 ng/ μ L | 4 | 10 ng/ μ L |
| Library compound | 1 mM | 0.2 | 10 μ M |
| SYPRO Orange | 8 x | 2.5 | 1 x |

*Reactions were prepared chronologically as listed

CHAPTER 3: RESULTS

3.1 Recombinant PafA expression, detection, and purification

3.1.1 Vector DNA isolation and characterisation

Following isolation and transformation of the constructs (pET24b(+)-PafA and pET24b(+)-PupEPafA) into their respective JM109 *E. coli* cells (DE3), the propagated cells were used to isolate the vector DNA and perform restriction digestion.

Undigested and single restriction digest reaction samples of vector DNA isolated from colonies of the pET24b(+)-PafA and pET24b(+)-PupEPafA JM109 (DE3) transformation reactions, are shown in Figure 3.1 A and B, respectively. The undigested samples both displayed the typical migration pattern observed for plasmid DNA i.e., nicked, and supercoiled conformations. Single digestion and linearisation of the pET24b(+)-PafA and pET24b(+)-PupEPafA plasmids showed extrapolated DNA sizes of 6.43 kbp and 6.31 kbp, respectively (APPENDIX Figure A1). These sizes correlated well with the expected size of the respective plasmid constructs, which were calculated using the known sizes of the pET24b(+) vector (5.31 kbp), the PafA (1.4 kbp) and pup (207 bp) genes. The pET24b(+)-PafA and pET24b(+)-PupEPafA had an expected calculated size of 6.6 kbp and 6.7 kbp, respectively.

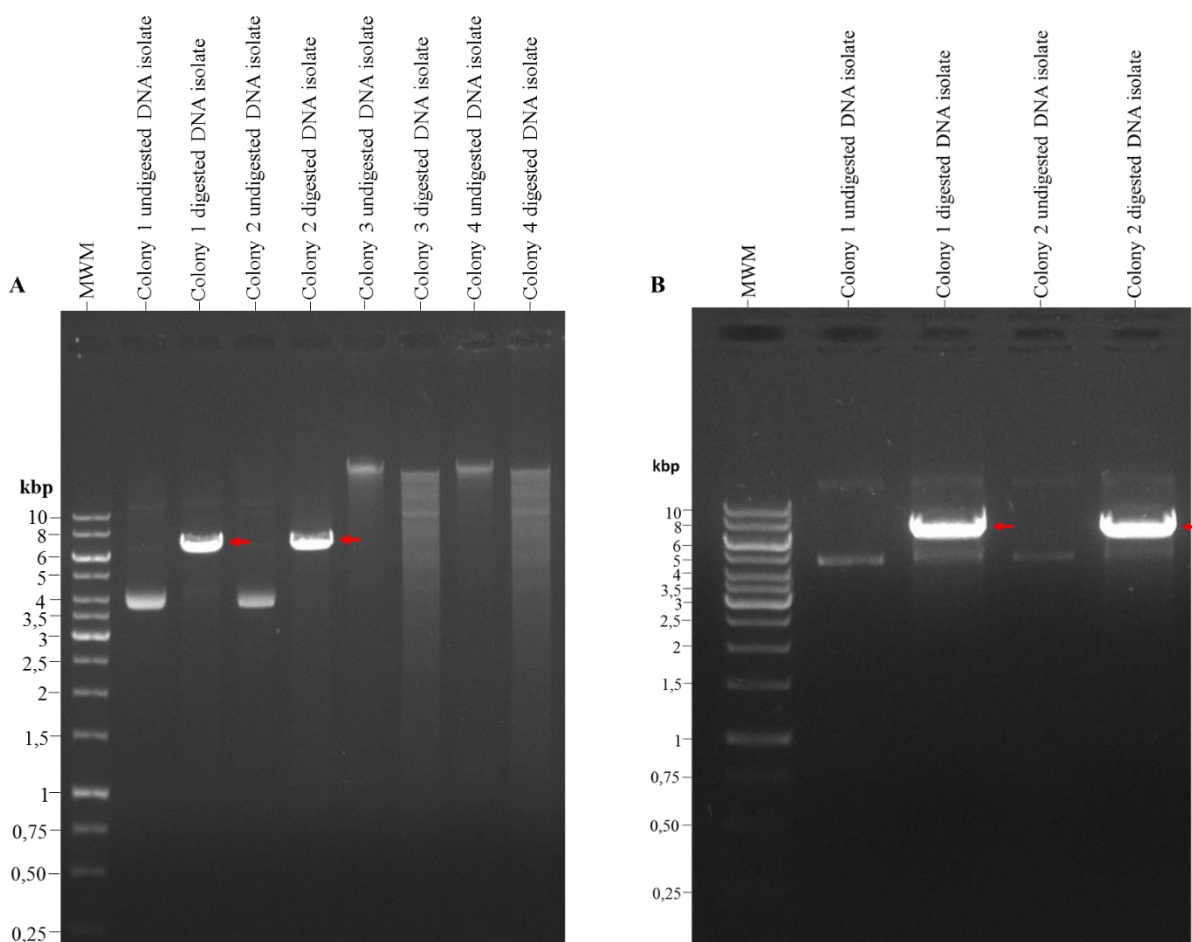


Figure 3.1. Agarose gel electrophoresis of pET24b(+)-PafA (A) and pET24b(+)-PupEPafA (B) recombinant *E. coli* cells. Vector DNA isolated from a successful colony culture was digested with 0.5 U *Nde*I, and subsequently analysed on a 1% (w/v) agarose gel stained with 0.005% (v/v) ethidium bromide (Promega, Wisconsin, USA), viewed under ultraviolet light. Lane MWM, O'GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific, Massachusetts, USA), and individual lanes from each gel showing either digested or undigested DNA. The arrows indicate vector DNA single restriction digest products.

After successful vector DNA isolation, PCR was conducted to detect the PafA and pup genes. Amplification of each gene product (Figure 3.2) was shown for both the pET24b(+)-PafA and pET24b(+)-PupEPafA vector DNA constructs with extrapolated sizes of 1.6 kbp and 204 bp, respectively. These extrapolated DNA sizes were consistent with the expected size of the relevant amplified gene i.e., PafA (1.4 kbp) and PupE (207 bp).

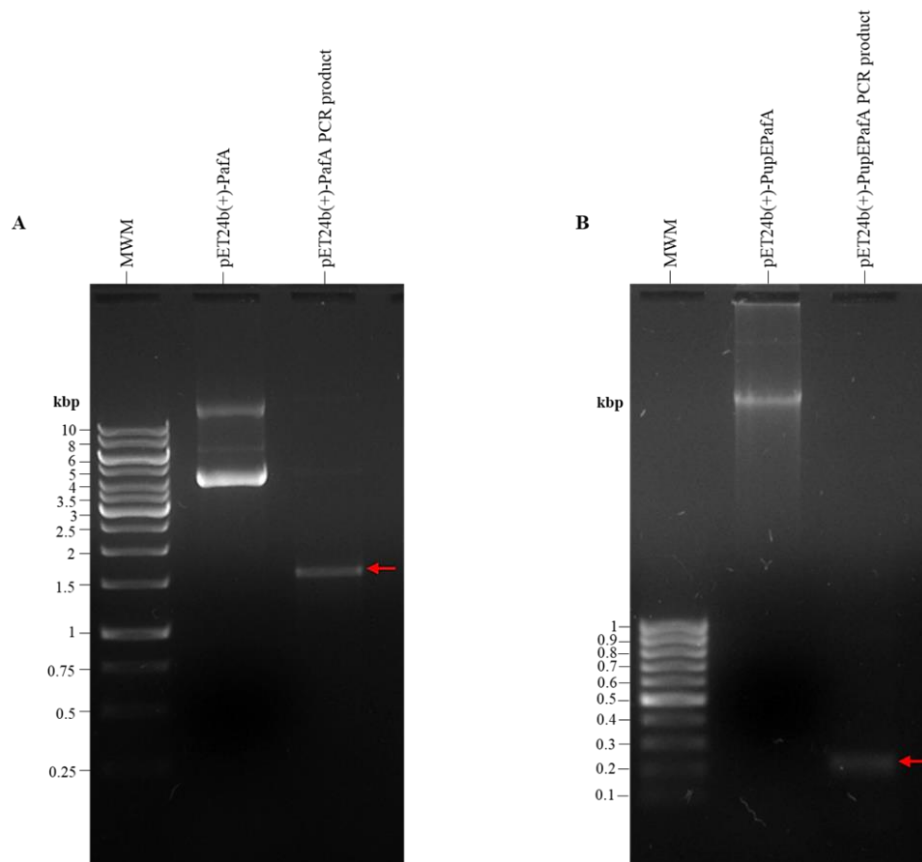


Figure 3.2. Amplification by PCR of the PafA (A) and pup (B) genes from the pET24b(+)-PafA and pET24b(+)-PupEPafA plasmids, respectively. After vector DNA (10 ng) was aseptically transferred into a PCR mixture of primers, 1 x *Thermus aquaticus* polymerase master mix and PCR conducted, the subsequent reactions were analysed on a 1% (w/v) agarose gel stained with 0.005% (v/v) ethidium bromide and viewed under ultraviolet light. Lane MWM, (A) O'GeneRuler 1 kbp DNA ladder and (B) GeneRuler 100 bp DNA ladder (Thermo Fisher Scientific, Massachusetts, USA). Lane 2 and lane 3 from each gel show recombinant vector DNA and PCR product migration, respectively. The arrows show the migrated PCR products for each plasmid construct.

Sequencing was also conducted to further confirm that the correct plasmid, containing either unmutated PafA and pup genes, was successfully isolated. Using the BLASTp NCBI database, PafA sequenced from the pET24b(+)-PafA construct returned a 100% parent identity alignment with *Mycobacterium tuberculosis* H37Rv PafA (accession number: NP_216613.1) and an E-value of 2×10^{-61} . PupE sequenced from the pET24b(+)-PupEPafA construct displayed significant similarity to pup from *Mtb* (accession number: QTR40672.1), reported with a 100% parent identity alignment and an E-value 7×10^{-26} . It was therefore concluded that the correct plasmids containing each gene of interest were isolated.

3.1.2 Recombinant PafA protein expression and immunoblot analysis

Following the transformation of JM109 (DE3) and BL21 (DE3) *E. coli* cells with either pET24b(+)-PafA or pET24b(+)-PupEPafA constructs, the transformed cells were induced with 1 mM IPTG to assess recombinant expression of PafA at 37 °C (Figure 3.3 and 3.4).

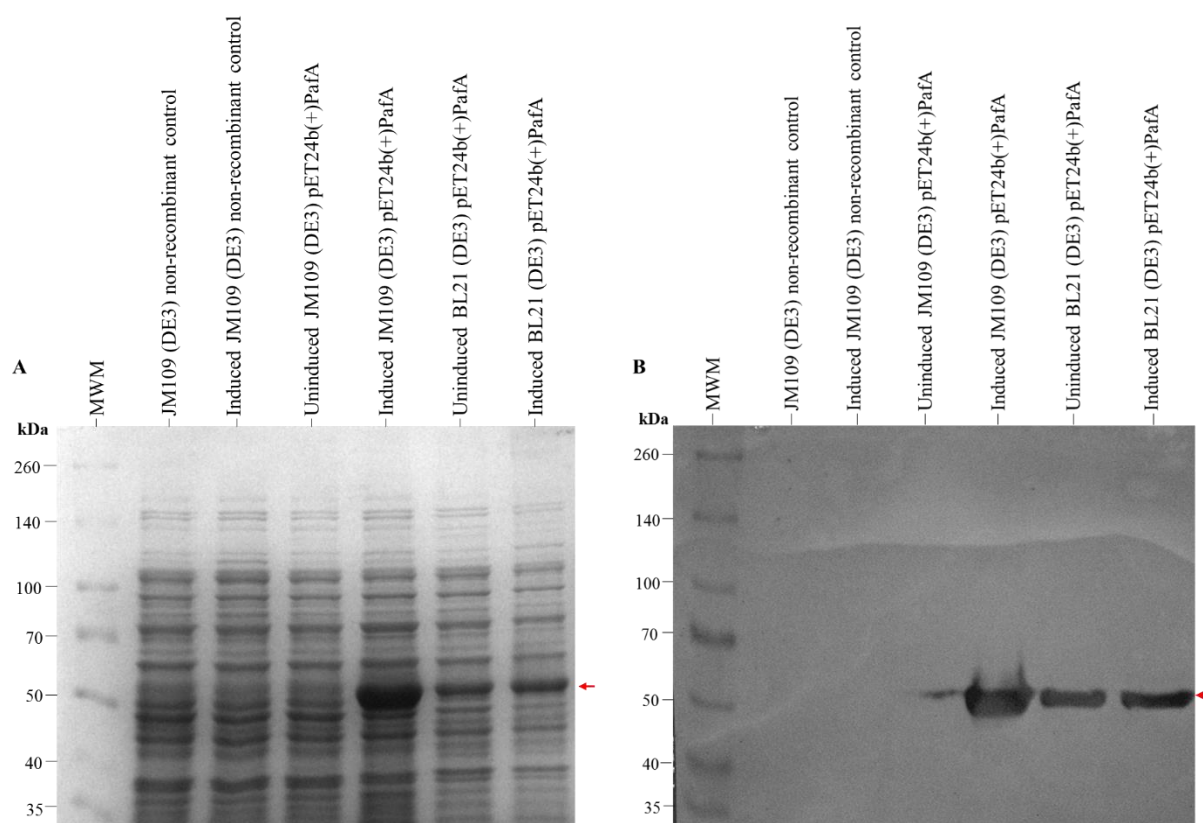


Figure 3.3. Recombinant JM109 (DE3) and BL21(DE3) expression of PafA, from pET24b(+)-PafA, as analysed by SDS-PAGE and immunoblot. Recombinant and non-recombinant *E. coli* cells were either left untreated or induced with 1 mM isopropyl β -d-1-thiogalactopyranoside (IPTG) for 4 h at 37 °C, and cell culture recovered by centrifugation. For analysis, culture samples were lysed (lysozyme (Thermo Fisher Scientific, Massachusetts, USA) and sonication) and samples analysed by (A) Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA) stained 10% reducing SDS-PAGE gel and (B) immunoblot analysis using mouse anti-6xHis 1° antibody (Catalog No.: R93025), Thermo Fisher Scientific, Massachusetts, USA), horse radish peroxidase (HRP)-conjugated anti-mouse 2° antibody (Catalog No.: 31430, Thermo Fisher Scientific Massachusetts, USA), and enhanced chemiluminescence (ECL) substrate (Novex, Massachusetts, USA). Lane MWM, Spectra Multicolour Broad Range Protein Ladder, and the arrows show migration of a 50-kDa protein.

The non-recombinant JM109 control, Figure 3.3A lane 2, presented the native whole-cell protein expression profile of the cell line. From the figure, no significant change in the protein profile was observed when the control non-recombinant JM109 *E. coli* cells were exposed to 1

mM IPTG (37 °C, 4 h) (lane 3). Similarly, the recombinant JM109 (DE3) *E. coli* cells in the absence of IPTG (lane 4) displayed a profile that was identical to the non-recombinant control. Exposing the pET24b(+)-PafA recombinant JM109 *E. coli* cells to 1 mM IPTG, Figure 3.3A lane 5, resulted in the expression of a 50-kDa protein (estimated using Rf values; APPENDIX Figure A2) not observed in the control JM109 *E. coli* cell samples. This protein size corresponded to the expected size of the PafA protein, reported in literature to be 51-kDa (Özcelik *et al*, 2012). In contrast to JM109 PafA protein expression, both the uninduced and induced pET24b(+)-PafA recombinant BL21(DE3) *E. coli* cell samples displayed expression of the 50-kDa protein (Figure 3.3A, lane 6 and 7, respectively).

Immunoblot analysis was also conducted to determine if the 50-kDa protein in the SDS-PAGE is the 6xHis tagged PafA protein. The detection of the 50-kDa protein in the 1 mM IPTG induced recombinant JM109 (DE3) *E. coli* cells was shown (Figure 3.3B, lane 5), while no detection in all the control JM109 *E. coli* cells was observed (Figure 3.3B, lanes 2, 3 and 4). PafA was also detected in the recombinant BL21 (DE3) *E. coli* cell samples, both in the absence and presence of IPTG, displaying an uncontrolled expression system (Figure 3.3B lane 6 and 7).

A similar analysis was subsequently conducted using the dual gene expressing vector, pET24b(+)-PupEPafA. The SDS-PAGE gel (Figure 3.4A) illustrates the whole-protein profile of recombinant JM109 *E. coli* cells transformed with pET24b(+)-PupEPafA. The control recombinant JM109 (DE3) *E. coli* cells grown in the absence of IPTG, and the recombinant JM109 (DE3) *E. coli* cells, treated with 1 mM IPTG, displayed a similar protein profile, showing expression of a protein with an estimate size of 51-kDa (APPENDIX Figure A2). IPTG presence resulted in an observable decrease in whole-protein concentration (lane 3) compared to the control recombinant JM109 *E. coli* cell sample (lane 2). The 51-kDa expressed protein, was comparable with the previously expressed PafA (Figure 3.3), suggesting that this was also PafA encoded by the dual gene expressing vector, pET24b(+)-PupEPafA. To confirm this, immunoblot analysis (Figure 3.4B) was done to detect the 6xHis tagged PafA protein showing detection at 51-kDa for both the control and IPTG exposed recombinant JM109 *E. coli* cell samples. These observations confirmed PafA expression by the pET24b(+)-PupEPafA recombinant JM109 *E. coli* cells. Additionally, PafA detection in both the control and the IPTG exposed recombinant JM109 (DE3) *E. coli* cells further validated pET24b(+)-PupEPafA

isolation and transformation due to the plasmids design allowing basal PafA expression after disruption of the lac-operon during its construction (Cerdeira-Maira *et al*, 2020).

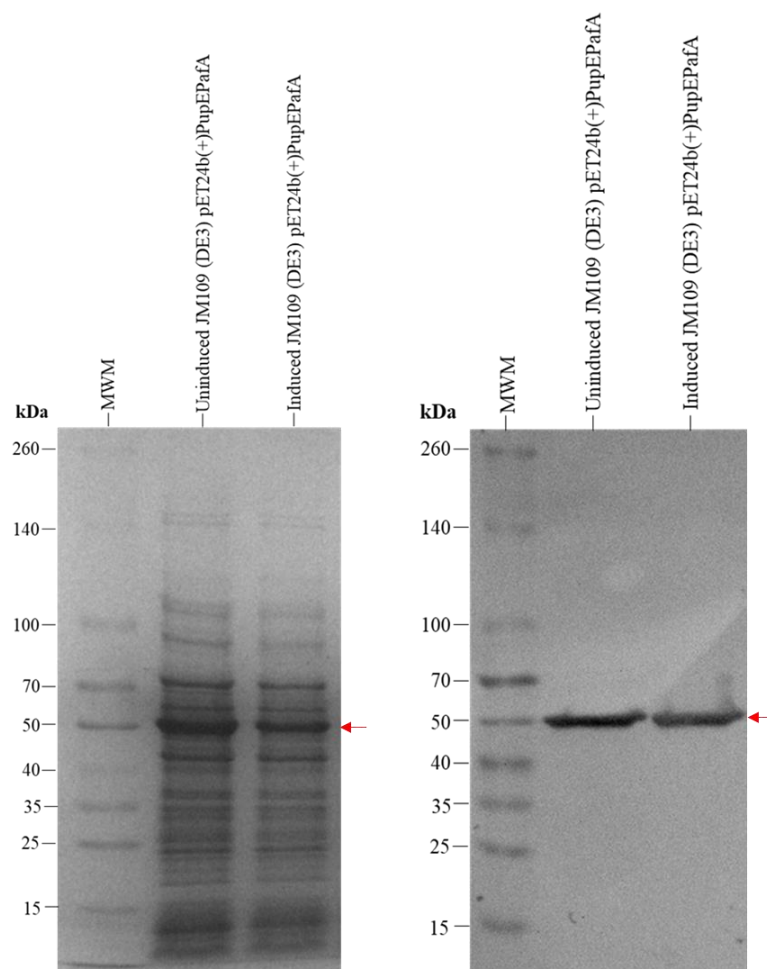


Figure 3.4. Recombinant JM109 (DE3) *E. coli* expression of PafA, from pET24b(+)-PupEPafA, as analysed by SDS-PAGE and immunoblot. Recombinant and non-recombinant *E. coli* cells were induced with 1 mM IPTG for 4 h at 37 °C, and cell culture recovered by centrifugation. For analysis, culture samples were lysed (lysozyme and sonication) and samples analysed by (A) Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA) stained 10% reducing SDS-PAGE gel and (B) immunoblot analysis using mouse anti-6xHis 1° antibody (Catalog No.: R93025), Thermo Fisher Scientific, Massachusetts, USA), horse radish peroxidase (HRP)-conjugated anti-mouse 2° antibody (Catalog No.: 31430, Thermo Fisher Scientific Massachusetts, USA), and enhanced chemiluminescence (ECL) substrate (Novex, Massachusetts, USA). Lane MWM, Spectra Multicolour Broad Range Protein Ladder, and the arrow shows migration of a 50-kDa protein.

Overall, PafA with a 6xHis tag was successfully expressed by both vectors. Since controllable expression of PafA was observed in pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cell samples (Figure 3.3), contrasting the uncontrolled PafA expression shown in recombinant

pET24b(+)-PafA BL21 (DE3) *E. coli* (Figure 3.3) and pET24b(+)-PupEPafA recombinant JM109 (DE3) *E. coli* cell samples (Figure 3.4), recombinant JM109 (DE3) cells transformed with the pET24b(+)-PafA plasmid construct were selected for further expression, optimisation, and protein purification.

3.1.3 PafA solubilisation and purification

PafA solubility was initially assessed in the pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cells at 37 °C for 4 h. Following detection of PafA expression exclusively in the insoluble fraction, PafA recombinant expression was performed at varying IPTG concentrations (Figure 3.5) and lower temperatures (Figure 3.6) in an attempt to obtain soluble PafA expression (Galloway *et al.*, 2003). SDS-PAGE analysis of pET24b(+)-PafA JM109 recombinant *E. coli* cells showed distinct expression of a 50-kDa protein exclusively in the insoluble fraction for all recombinant expressions performed at 37 °C, suggesting the presence of inclusion bodies (Figure 3.5)

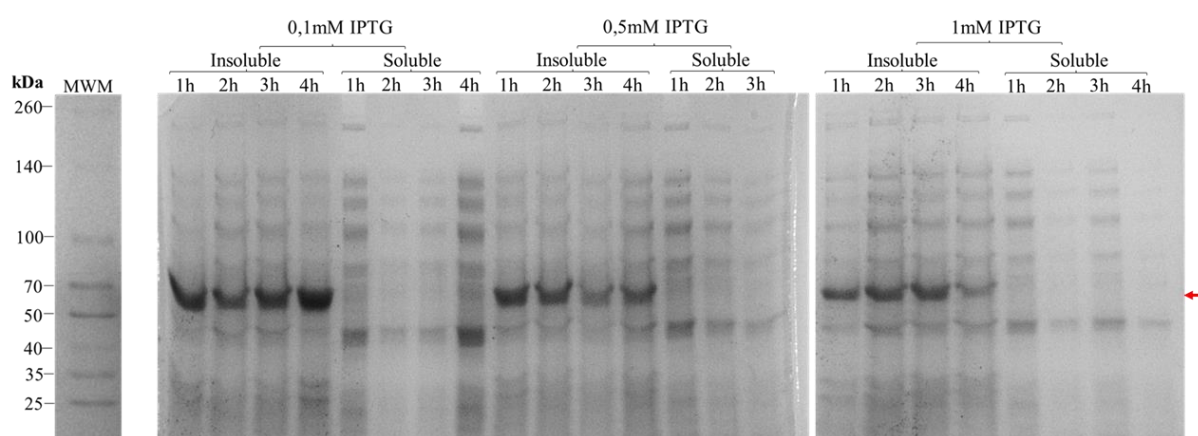


Figure 3.5. The solubility of PafA when IPTG concentrations were varied (4 h, 37 °C). Expression was initiated with IPTG (0.1 mM, 0.5 mM or 1.0 mM) and samples taken hourly for 4 h. Each expression sample was sonicated 15 x on ice, centrifuged at 12 000 x g for 30 min and soluble and insoluble fractions analysed via a 10% reducing SDS-PAGE stained with Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA). Lane MWM, a representative Spectra Multicolour Broad Range Protein Ladder. The arrow shows migration of a 50-kDa protein.

Expression at 20 °C for 20 h with 0.1 mM, 0.5 mM and 1.0 mM IPTG was therefore attempted to solubilise the protein (Figure 3.6), however, PafA remained in the insoluble fraction as inclusion bodies (Figure 3.6; lanes 2, 4, and 6). Due to the higher quality of recombinant

proteins typically recovered from inclusion bodies formed at low temperature (Jevševar et al., 2005), expression at 20 °C for 20 h with 1 mM IPTG was favoured and selected for all subsequent expression of PafA.

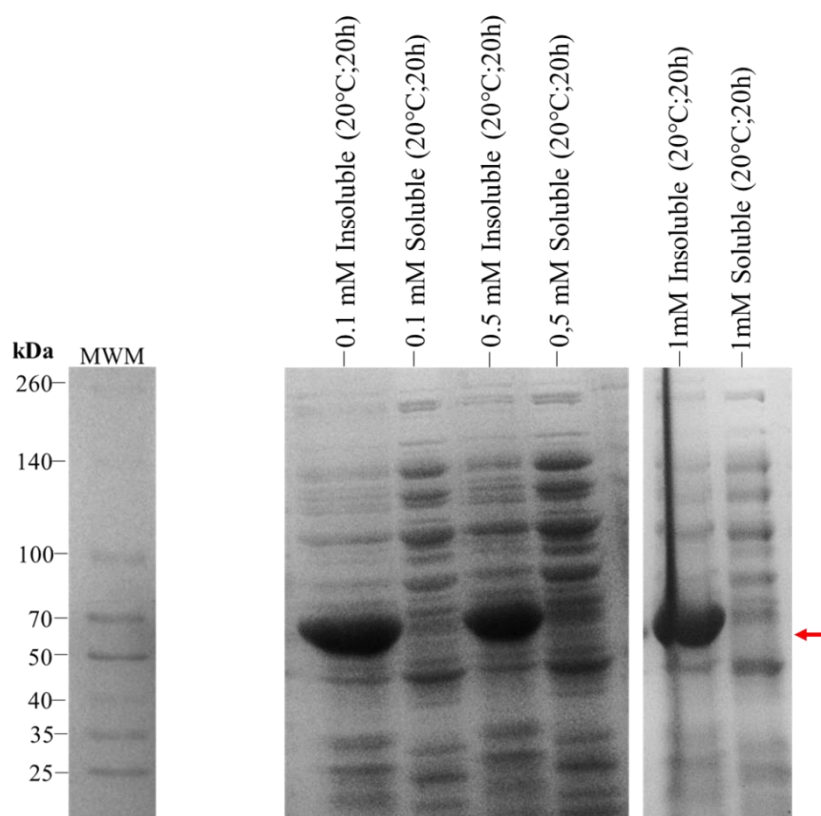


Figure 3.6. PafA solubility at low temperature expressions (20 h, 20 °C). Recombinant and non-recombinant *E. coli* cells were induced with 1 mM IPTG for 4 h at 37 °C, and cell culture recovered by centrifugation (5 000 x g, 10 mi at 4 °C). For analysis, culture samples were lysed (750 µg/mL lysozyme and 15 x sonication at 50% power (30-s sonication, 30-s off)) and samples analysed by via a 10% reducing SDS-PAGE gel stained with Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA). Lane MWM, a representative Spectra Multicolour Broad Range Protein Ladder, and the arrow shows the migration of a solubilised 50-kDa protein.

The insoluble expression of PafA (Figure 3.5 and 3.6) prompted the need to solubilise the protein for subsequent IMAC purification (Figure 3.7). SDS (1% w/v) detergent was selected to solubilise aggregated PafA due to compatibility with the TALON IMAC resin and easy removal by low temperature precipitation (Schlager *et al*, 2012). As shown previously, PafA recombinant expression in pET24b(+)-PafA transformed JM109 (DE3) *E. coli* cells was found exclusively in the insoluble fraction as inclusion aggregates (Figure 3.7, lane 2; Figure 3.8 lane 2). However, the addition of 1% (w/v) SDS resulted in an increase of PafA in the soluble

fraction (Figure 3.7, lane 5) and a decrease in the insoluble fraction (Figure 3.7 lane 4) when compared to the untreated (no detergent) control samples (Figure 3.7 lanes 1 and 2).

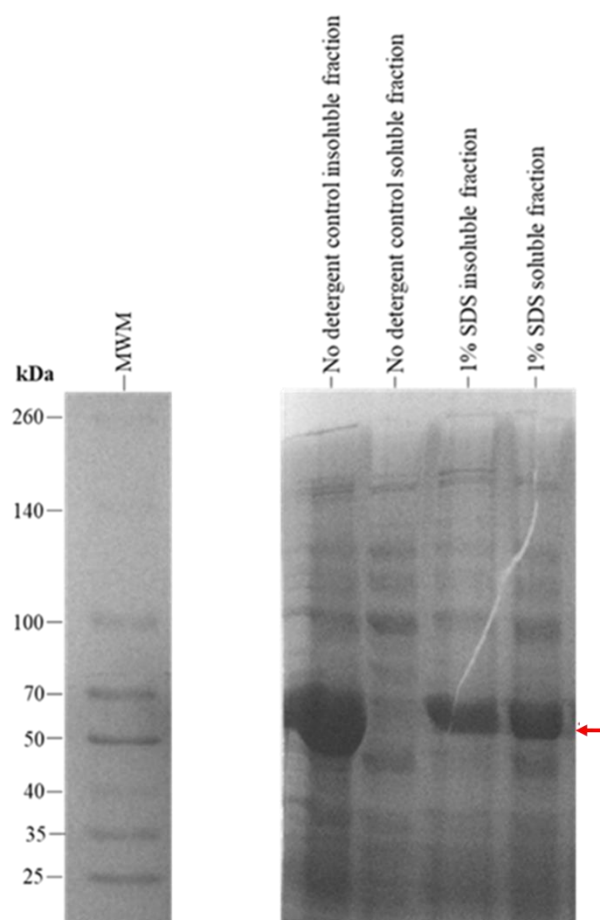


Figure 3.7. SDS solubilisation of PafA from inclusion bodies. Pelleted expression culture (5000 x g for 10 min, 4 °C) was lysed with 750 µg/mL lysozyme, sonicated 15 x on ice (50% power; 30-s sonication, 30-s off) and the soluble and insoluble fractions collected by centrifugation (12 000 x g for 30 min, 4 °C). The pellet was resuspended using 1% (w/v) SDS in 1xEB (1 x equilibration buffer; 50 mM Na₂HPO₄, 300 mM NaCl), sonicated, centrifuged, and the soluble and insoluble fractions separated on a 10% reducing SDS-PAGE gel which was stained with Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA). Lane MWM, a representative Spectra Multicolour Broad Range Protein Ladder. The arrow indicates the migration of a solubilised 50-kDa protein.

The solubilised PafA protein sample was subsequently used in IMAC purification (Figure 3.8). Lanes 2 and 3 from Figure 3.8A showed the soluble and insoluble protein profiles of pET24b(+)-PafA IPTG induced *E. coli* cells lysed with lysozyme (750 µg/mL) and sonicator (15 x at 50% power; 30s sonication, 30s rest), respectively. Figure 3.8B lane 2 and lane 3 represented immunoblot analysis of the same samples, detecting PafA expression exclusively in the insoluble fraction. Upon the addition of 1% SDS, Figure 3.8A lane 6, a considerable increase of PafA in the soluble fraction was observed, an increase that was also detected by

immunoblot analysis shown by Figure 3.8B lane 6. After exposure to the TALON resin for 1 h at 4 °C, an unbound sample was collected (lane 7 Figure 3.8A and Figure 3.8B). A substantial decrease in PafA concentration was seen in the unbound sample compared to the solubilised PafA sample prior to resin application (Figure 3.8A lane 7). This observation was also detected by immunoblot analysis (Figure 3.8B lane 7), suggesting PafA successfully bound to the resin. After washing the resin bed 5 x with 5 mM imidazole 1xEB buffer (Figure 3.8A and Figure 3.8B, lanes 8 to 10), PafA was successfully recovered from the resin bed with 250 mM imidazole as observed by SDS-PAGE as a single 50-kDa protein and confirmed by immunoblot analysis (Figure 3.8A and B, lane 11 to lane 13).

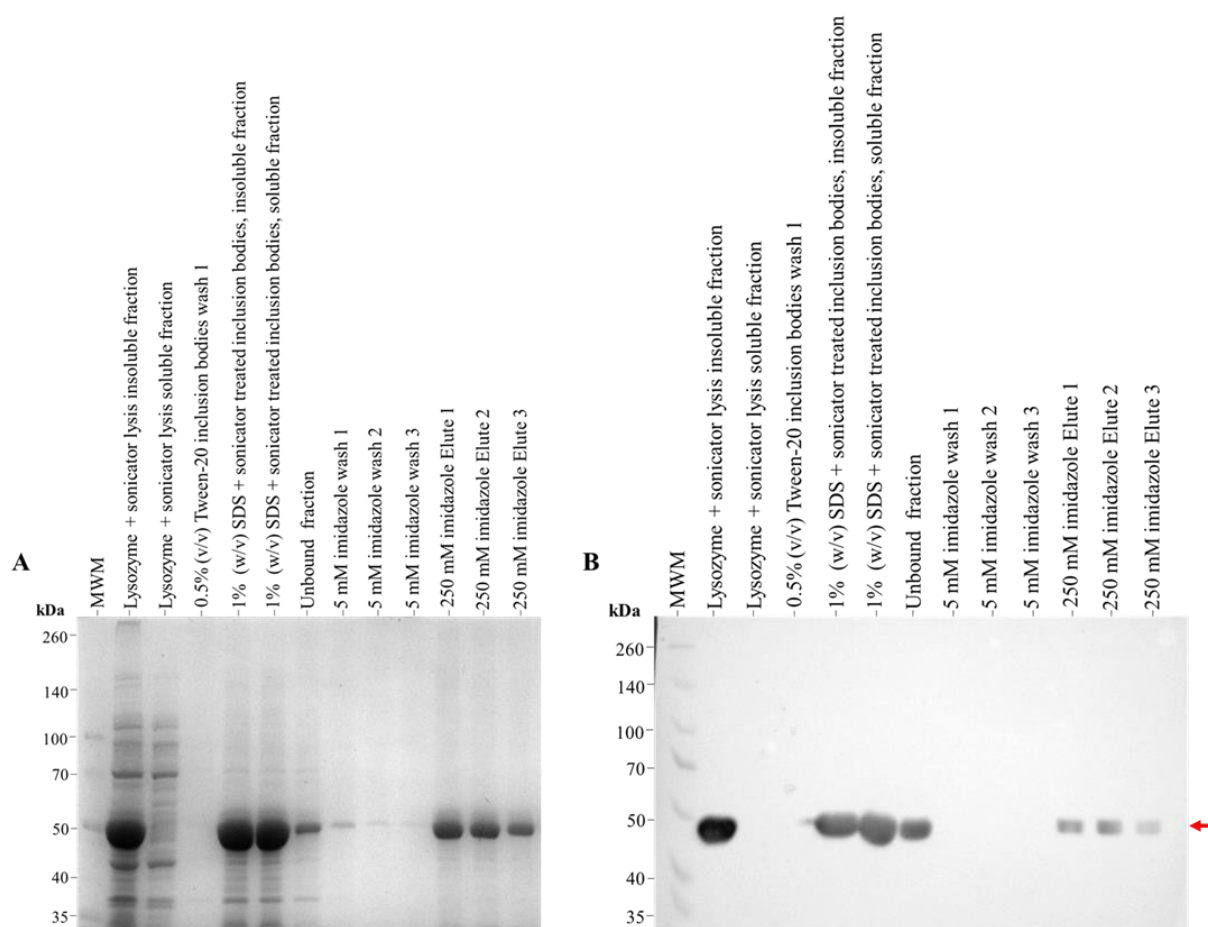


Figure 3.8. PafA IMAC purification using TALON resin. Recombinant PafA was expressed (1 mM IPTG, 20 °C, 20 h), lysed with 750 µg/mL lysozyme (37 °C, 30 min) and sonication (15 x on ice at 50% power; 30-s sonication, 30-s rest), and the PafA aggregates recovered by centrifugation (12 000 x g for 30 min, 4 °C). The pellet fraction was washed 2 x with 0.5% (v/v) Tween-20 in 1xEB, and PafA aggregates solubilised using 1% (w/v) SDS in 1xEB and sonication. The sample was centrifuged, and the soluble fraction subsequently exposed to a pre-equilibrated TALON resin. After 1 h incubation at 4 °C, the resin was washed with 5 mM imidazole in 1xEB, and bound protein eluted with 250 mM imidazole in 1xEB. Samples were analysed by 10% reducing SDS-PAGE stained with (A) Bromophenol blue (Thermo Fisher Scientific, Massachusetts, USA) and (B) immunoblot using mouse anti-6xHis 1° antibody (Catalog No.: R93025), Thermo Fisher Scientific, Massachusetts, USA), horse radish peroxidase (HRP)-conjugated anti-mouse 2° antibody (Catalog No.: 31430, Thermo Fisher Scientific Massachusetts, USA), and enhanced chemiluminescence (ECL) substrate (Novex, Massachusetts, USA). Lane MWM, Spectra Multicolour Broad Range Protein Ladder.

3.2. The detection of aggregating PAINs via a colorimetric enzyme assay

3.2.1 Enzyme assay

The enzyme HRP and its chromogenic substrate TMB, were used to develop an assay that detects aggregating PAIN compounds. Initially, the colourimetric assay conditions were optimised by varying HRP enzyme concentration (5 - 100 ng/mL in PBS, pH 7.4) against a constant TMB concentration (50 ng/mL). Data from the study, shown in Figure 3.9, was used to determine the optimal HRP concentration for the colourimetric assay. An increase in HRP concentration from 5 ng/mL to 100 ng/mL resulted in a significant increase in the overall colour formation ($p < 0.0001$), detecting more product formation at higher enzyme concentrations, as anticipated. Consequently, a 100 ng/mL HRP concentration was selected as the appropriate concentration for the standard enzymatic assay, displaying acceptable levels of detection ($A_{652\text{nm}} = 0.71 \pm 0.016$) at a relatively low HRP concentration.

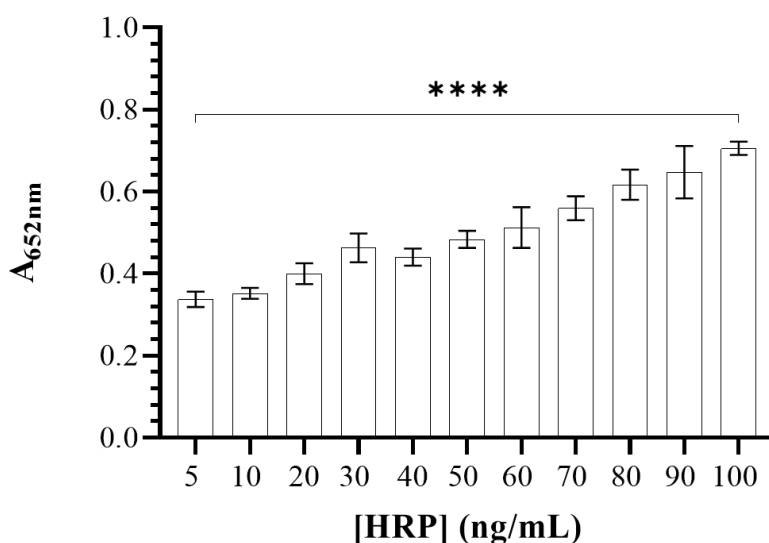


Figure 3.9. The effect of HRP enzyme concentration on colourimetric development in the presence of TMB substrate. Increasing HRP concentration (5 – 100 ng/mL) were prepared in PBS buffer (pH 7.4) and the enzymatic reaction initiated by the addition of 50 ng/mL TMB substrate solution (0.02% H_2O_2). Absorbance reading, as mean \pm standard deviation at 652 nm, after 10 min incubation at room temperature are shown ($n = 3$). An unpaired Welch's t-test was done to determine significance of HRP concentrations displaying $> 0.7 A_{652\text{nm}}$ against the initial 5 ng/mL HRP, where **** $p < 0.0001$

3.2.2 Non-specific inhibitor screening assay

A selection of four compounds, previously described to possess non-specific inhibitor properties, were used to determine their effect on the enzyme activity of the colourimetric assay. These included quercetin, EGCG, tannic acid, and 8-hydroxyquinoline. Sodium azide, previously reported as an irreversible HRP inhibitor ($K_i = 1.47$ mM; K_i , inhibitor constant i.e., compound concentration for half maximum inhibition) (Ortiz de Montellano *et al.*, 1988), was included as the positive control for the assay, contrasted by the two recruited negative controls caffeine and warfarin. Another plant-derived compound, riboflavin (vitamin B2), was included in the assay to test the sensitivity of the assay against pigmented compounds. Caffeine and warfarin, as expected, did not significantly affect the colorimetric assay when compared to the no compound HRP control ($p = 0.1033$ and $p = 1278$, respectively). Similarly, Riboflavin did not significantly affect HRP activity when compared to the no compound control ($p = 0.0945$). EGCG, quercetin, and tannic acid displayed the most potent inhibiting effect on HRP activity, with all three resulting in > 95% HRP activity inhibition (Figure 3.10). Significant inhibiting activity was observed for 8-hydroxyquinoline, resulting in 76.1% reduction of HRP activity ($p < 0.01$). Riboflavin, however, did not significantly affect HRP activity when compared to the no compound control ($p = 0.0945$).

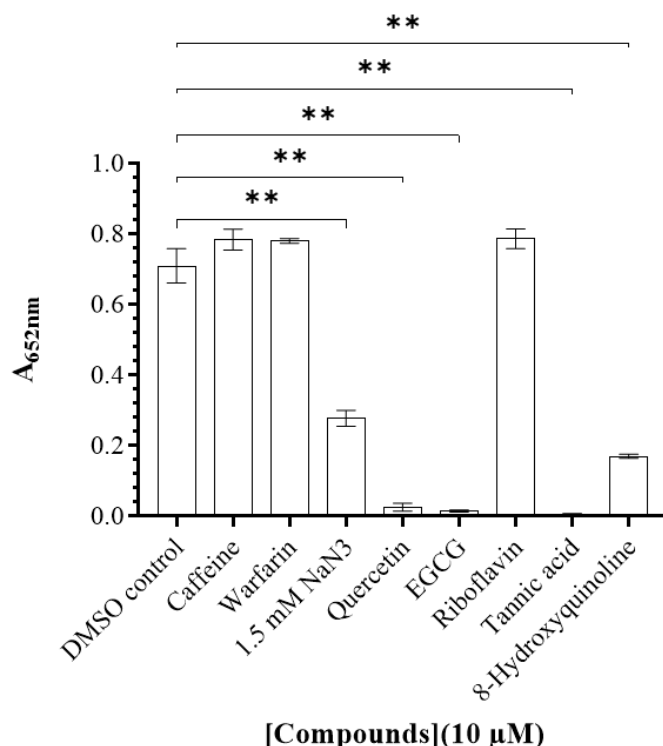


Figure 3.10. The effect of non-specific on the colourimetric HRP enzyme assay. Compounds (10 µM) were pre-incubated with 100 ng/mL HRP (PBS, pH 7.4) for 5 min, and the enzymatic reaction initiated by the addition of 50 ng/mL TMB substrate solution (with 0.02% H₂O₂). Absorbance readings, as mean ± standard deviation at 652 nm after 10 min incubation at room temperature (n = 3), are shown and an unpaired Welch's t-test done to determine the significant effect of each compound on the HRP enzymatic assay, where ** $p < 0.01$.

Next, the developed non-specific screening assay was utilised to screen for potential non-specific inhibiting compounds from the DiverSET CHEMBRIDGE small-molecule compound library. None of the DiverSET CHEMBRIDGE library compounds assayed at 10 µM, significantly affected the HRP-TMB enzymatic assay i.e., $p > 0.05$ for all assayed compounds compared to the DMSO control (Figure 3.11). This was in line with the virtual screening performed for the compound library on the ZINC15 chemical database which predicted that all the 48 DiverSET CHEMBRIDGE compounds lack aggregation inhibiting activity based on their chemical structure. The observed 62.4% HRP inhibition by compound C11 was determined to be statistically insignificant due to the high variability observed in its data ($p = 0.2210$). As observed previously, and included here for reference, the two negative controls caffeine and warfarin did not result in any significant effect on the colorimetric assay ($p = 0.1995$ and $p = 0.2206$, respectively) contrasting the positive control NaN₃ ($p \leq 0.001$)

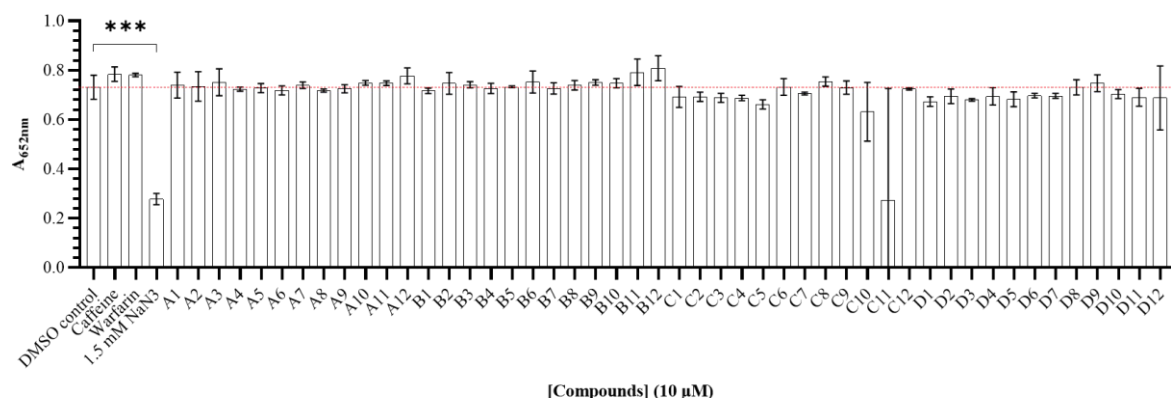


Figure 3.11 Screening 48 compounds from the DiverSET CHEMBRIDGE library for non-specific inhibition. Each compound (10 μ M) was pre-incubated with 100 ng/mL HRP (PBS, pH 7.4) for 5 min, and the enzymatic reaction initiated by the addition of 50 ng/mL TMB substrate solution (with 0.02% H_2O_2). Caffeine, warfarine, and NaN_3 were included as reference controls for the assays. Absorbance readings, as mean \pm standard deviation at 652 nm after 10 min incubation at room temperature ($n = 3$), are shown and an unpaired Welch's t-test done to determine the significant effect of each compound on the HRP enzymatic assay, where *** $p < 0.001$.

3.3 Screening of small molecules for PafA-binding capability via a thermal shift assay (TSA)

Prior to optimising a TSA for PafA, a control thermal melt reaction was carried out using 100 ng/ μ L control thermal shift protein, 1 mM control thermal shift ligand, and 5 x SYPRO Orange fluorescence dye. Thermal melt analysis of the control protein against 5 x SYPRO Orange fluorescence dye revealed this protein to have a T_m of 40.9 ± 0.5 $^{\circ}\text{C}$ (Figure 3.12B). Upon the addition of the control ligand, the T_m of the control protein significantly increased by 5.7 $^{\circ}\text{C}$ ($p < 0.05$) resulting in a new T_m of 46.6 ± 0.1 $^{\circ}\text{C}$ (Figure 3.12B).

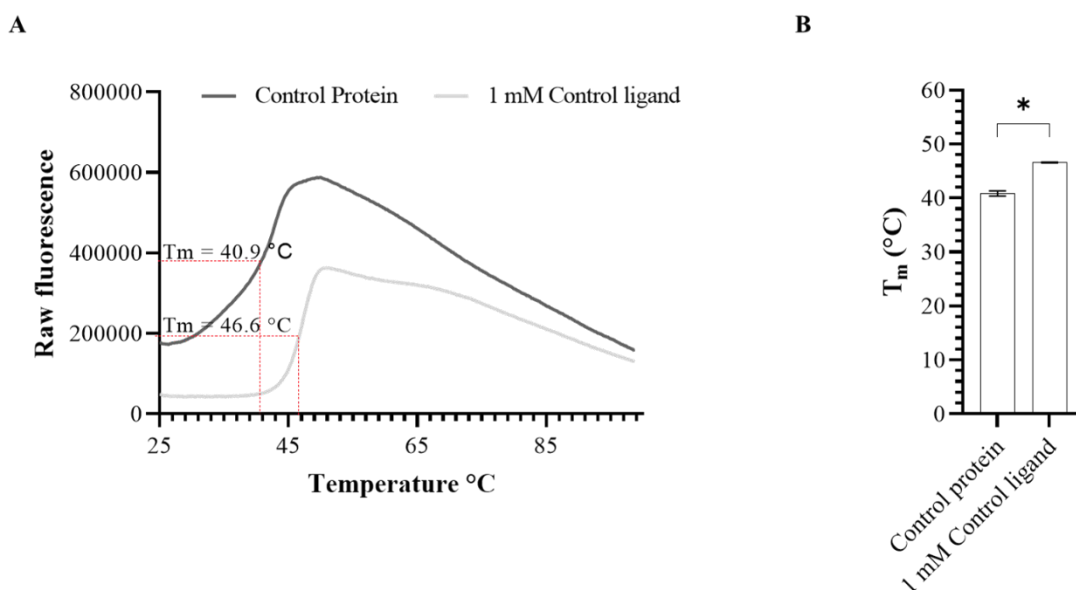


Figure 3.12. Control thermal shift assay using the protein thermal shift starter kit. The control protein (100 ng/ μL) was pre-incubated in a PBS-MilliQ deionized water solution in the presence or absence of 1 mM control ligand for 30 min on ice. SYPRO Orange dye stock solution was added to a 5 x final concentration and the solution mixed, centrifuged (10 300 x g, 10-s), and placed in the dark on ice until thermal melt examination. Analysis was done from 25 °C to 99 °C at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths where (A) 100 ng/ μL control protein thermal melt curves in the presence or absence of 1 mM control ligand ($R^2 > 0.99$), and (B) the extrapolated melting temperatures (T_m), as mean \pm standard deviation, of the control protein. The red lines in (A) show the control protein T_m in the presence and absence of the control protein ligand, extrapolated from its respective thermal melt reactions and taken as the EC_{50_1} after non-linear regression analysis and * $p < 0.05$.

Next, dialysed pure PafA and SYPRO Orange were used to design and optimise a TSA that can be used to screen for possible PafA-binding ligands from the DiverSET CHEMBRIDGE small molecule compound library.

To determine the best conditions for the TSA, increasing PafA concentration (1 ng/ μL , 2.5 ng/ μL , 5 ng/ μL , 7.5 ng/ μL , 10 ng/ μL , 50 ng/ μL , 100 ng/ μL , 150 ng/ μL , 200 ng/ μL , 250 ng/ μL , and 300 ng/ μL) was assayed against increasing SYPRO Orange concentrations (1 x, 5 x, 10 x, 15 x, and 20 x). It was observed that 10 ng/ μL PafA against 1 x SYPRO Orange dye was the optimal conditions, shown in Figure 3.13A, where PafA T_m was determined to be $47.8 \pm 1.2^\circ\text{C}$ (Figure 3.13B).

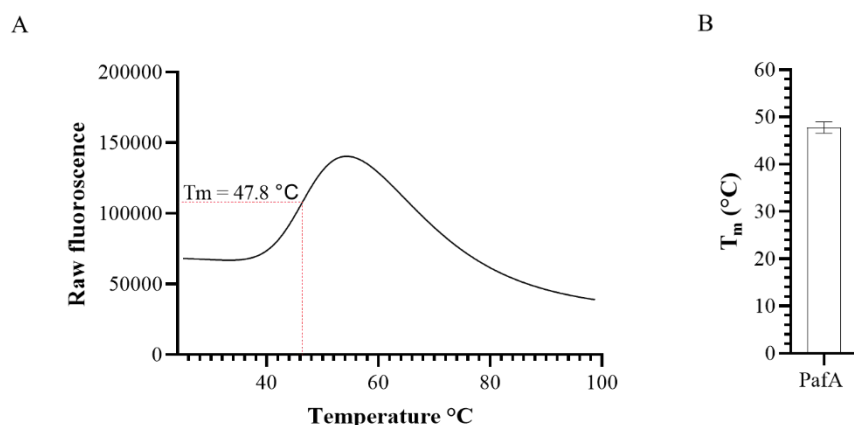


Figure 3.13. Optimised PafA thermal shift assay. PafA (10 ng / μ L) was pre-incubated in a PBS-MilliQ deionized water solution for 30 min on ice. SYPRO Orange dye solution was added to 1 x final concentration and the solution mixed, centrifuged, and placed in the dark on ice until thermal melt examination. Analysis was done from 25 °C to 99 °C at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths where (A) shows 10 ng/ μ L PafA thermal melt against 1 x SYPRO Orange dye and ($R^2 = 0.98$) (B) mean PafA melting temperature (T_m) \pm standard deviation ($n = 2$). The red line in (A) shows PafA T_m extrapolation from its thermal melt reaction, taken as the EC_{50_1} after non-linear regression analysis.

3.3.1 PafA-binding ligand screening via the optimised PafA TSA

Once the TSA was optimised for PafA, the 48-compounds from the DiverSET CHEMBRIDGE library previously subjected to the non-specific inhibitor detecting assay were screened for PafA-binding efficacy as shown in Figure 3.14.

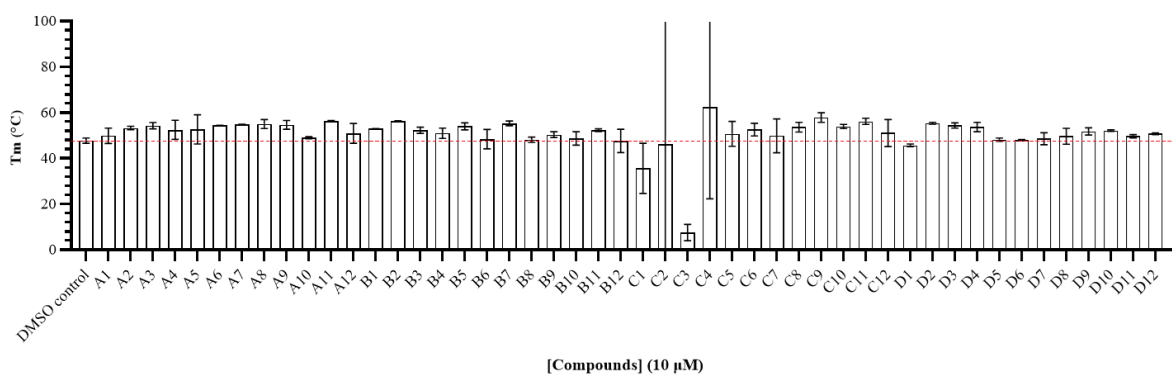


Figure 3.14. Screening PafA for possible binding ligands using the TSA. PafA (10 ng / μ L) was preincubated in PBS-MilliQ deionized water in the presence or absence of 10 μ M DiverSET CHEMBRIDGE compound for 30 min on ice. SYPRO Orange dye stock solution was subsequently added to a final 1 x concentration and the solution gently mixed, centrifuged, and PafA thermal melt assessed from 25 °C to 99 °C at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths. T_m readings, as mean \pm standard deviation extrapolated from the respective thermal melt reactions and taken as the EC_{50_1} after non-linear regression analysis, are shown ($n = 2$) and the red horizontal line drawn to highlight PafA T_m increase in the presence of the DiverSET CHEMBRIDGE compounds.

Figure 3.14 shows PafA T_m in the presence of 10 μ M DiverSET CHEMBRIDGE compounds. PafA T_m in the presence of the individual compounds was compared to the control DMSO PafA T_m, where compounds displaying a significant effect on PafA T_m i.e., ≥ 2 °C increase in PafA T_m and $P < 0.05$, were selected as appropriate potential PafA binders. Table 3.1 demonstrates the DiverSET CHEMBRIDGE compounds that induced a significant shift in PafA T_m at 10 μ M, ranked according to greatest change in T_m and lowest p value.

Table 3.1. DiverSET CHEMBRIDGE compounds that significantly increased PafA T_m

| Compound | Melting T _m (°C) | T _m shift (°C) | P-value | R ² -value | Significance ($P < 0.05$) |
|----------|--------------------------------|------------------------------|---------|-----------------------|--------------------------------|
| C9 | 57.8 \pm 2.1 | 10.1 | 0.0451 | 0.9568 | Yes |
| C11 | 56.2 \pm 1.3 | 8.4 | 0.0224 | 0.9575 | Yes |
| B7 | 55.3 \pm 1.1 | 7.5 | 0.0227 | 0.9571 | Yes |
| D3 | 55.4 \pm 1.2 | 7.6 | 0.0296 | 0.9420 | Yes |
| D2 | 55.4 \pm 0.4 | 7.6 | 0.0484 | 0.9835 | Yes |
| A3 | 54.2 \pm 1.4 | 6.4 | 0.0375 | 0.9295 | Yes |
| C10 | 53.9 \pm 0.9 | 6.1 | 0.0330 | 0.9483 | Yes |
| A2 | 53.0 \pm 0.4 | 5.2 | 0.0436 | 0.9452 | Yes |

An unpaired Welch's t-test was done to determine the significant effect of each compound on PafA T_m (T_m; temperature where 50% of a protein is unfolded.)

Compound C9 yielded the highest shift in T_m relative to the control PafA T_m, resulting in an increased shift of 10.1 °C from 47.8 °C \pm 1.2 °C to a T_m of 57.8 \pm 2.1 °C (Table 3.1). Similarly, compounds C11 (T_m = 56.2 \pm 1.3 °C); B7 (T_m = 55.3 \pm 1.1 °C); D3 (T_m = 55.4 \pm 1.2 °C); D2 (T_m = 55.4 \pm 0.4 °C); A3 (T_m = 54.2 \pm 1.4 °C); C10 (T_m = 53.9 \pm 0.9 °C); and A2 (T_m = 56.3 \pm 0.2 °C), also induced a significant (≥ 2 °C PafA T_m increase; $P < 0.05$) shift in PafA T_m, each compound causing a shift of 8.4 °C; 7.5 °C; 7.6 °C; 7.6 °C; 6.4 °C; and 6.1 °C, respectively (Table 3.1). Chemical structures of the compounds that induced a significant shift in PafA T_m are shown in Figure 3.15.

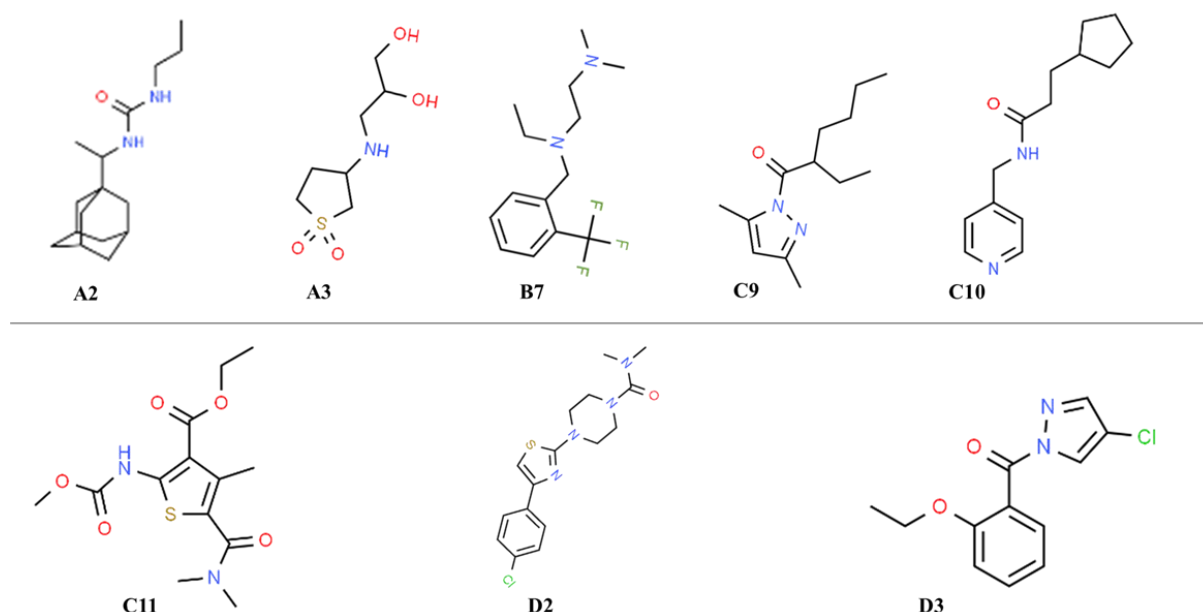


Figure 3.15. Molecular structure of the potential PafA binders. The chemical structures were sourced from PubChem, an online chemical database (<https://pubchem.ncbi.nlm.nih.gov/>)

Compounds A8 ($T_m = 55 \pm 2.0$ °C); A7 ($T_m = 54.9 \pm 0.1$ °C); A9 ($T_m = 54.6 \pm 1.9$ °C); A6 ($T_m = 54.5 \pm 0.1$ °C); A3 ($T_m = 54.2 \pm 1.4$ °C); and D4 ($T_m = 53.7 \pm 2.1$ °C), also considerably effected PafA T_m , resulting in an increased T_m shift of 7.2 °C; 7.1 °C; 6.8 °C; 6.7 °C; 6.4 °C; and 5.9 °C, respectively (Appendix Table A1). Data from these compounds, however, was statistically insignificant against the control DMSO PafA T_m , with all the compounds resulting in P -values > 0.05 . Similarly, other compounds that observed a notable shift in PafA T_m - albeit with high data variation ($P > 0.05$), included C4 ($T_m = 62.4 \pm 40.10$ °C); C12 ($T_m = 51.1 \pm 5.89$ °C); A12 ($T_m = 50.9 \pm 4.3$ °C); C5 ($T_m = 50.7 \pm 5.4$ °C); A1 ($T_m = 49.9 \pm 3.39$ °C); and C7 ($T_m = 49.9 \pm 7.39$ °C), each with a perceived shift of 14.6 °C; 3.3 °C; 3.1 °C; 2.9 °C; 2.1 °C; and 2.1 °C respectively. APPENDIX Table A1 showcases the effect of all the DiverSET CHEMBRIDGE library compounds on PafA T_m .

Library compounds that either resulted in a negative T_m shift i.e., observed a lower T_m relative to the DMSO control, or observed no significant shift in PafA T_m (i.e., < 2 °C shift and/or $P > 0.05$) were eliminated as potential PafA-binding ligands. This included compounds A1; A4 - A12; B1 - B6; B8 - B12; C1 - C2; C4 - C8; C12; D1; and D4 - C12 (APPENDIX Table A1).

Taken together, the 8 compounds that induced a significant increase in PafA T_m were identified as possible binding ligands. The high number of compounds discovered to incur a significant increase in PafA T_m was unexpected. Therefore, further investigation to confirm these results is still required using alternative binding assays such as circular dichroism.

CHAPTER 4: DISCUSSION

The increasing emergence of *Mtb* strains resistant to the current first-line anti-TB drugs compromises global health security. Consequently, innovative anti-TB strategies are needed to circumvent resistant issues and help eradicate *Mtb* prevalence. A novel approach would be to recruit PROTAC technology as an alternative anti-TB treatment option by developing new PROTAC-like drugs capable of harnessing the *Mtb* protein recycling cascade system i.e., pupylation, to target *Mtb* essential proteins for apoptosis. To this effect, the current study looked to identify ligands capable of efficiently binding to the sole pupylating ligase in *Mtb*, PafA, as an essential first step in the development of the PROTAC-like prokaryote degrading drug(s).

4.1 *PafA* and *PupE* bacterial transformation

JM109 (DE3) *E. coli* cells were transformed for PafA recombinant expression using two plasmid constructs pET24b(+)-PafA and pET24b(+)-PupEPafA, both encoding the PafA gene. Growth on Kanamycin antibiotic medium was used as an initial indicator of successful transformation, both plasmid constructs housing the Kanamycin resistance gene in contrast to the control non-recombinant JM109 (DE3) *E. coli* cells. As anticipated, the control non-recombinant JM109 (DE3) *E. coli* cells did not develop colony growth on the antibiotic supplemented medium, which further implied that the observed survival and proliferation of cells as a colony forming unit from each respective transformation reaction to be dependent on the Kanamycin resistance gene conferred by the vector constructs. To further validate successful transformed with the two plasmids, each colony was propagated and the resulting culture medium used to isolate vector DNA i.e., pET24b(+)-PafA and pET24b(+)-PupEPafA constructs, for targeted single restriction digestion. Two out of the four colonies from the pET24b(+)-PafA transformation product were successfully digested using the type II restriction endonuclease *NdeI*, yielding a predominantly single band of linearised DNA sized 6.43 kbp. This size correlated with the calculated expected size of 6.6 kbp using SnapGene 5.2.3 software, virtually assembling pET24b(+)-PafA following the protocol used in its construction for PafA overexpression by Festa *et al.*, (2007). Similarly, single restriction digestion of the two colonies from the pET24b(+)-PupEPafA transformation products by *NdeI* resulted in a predominantly single band of linearised DNA sized 6.31 kbp, which also correlated with the calculated expected size of 6.7 kbp also virtually rendered using SnapGene 5.2.3 . The specific single digestion of both isolated vector DNA further emphasised the successful transformation of the

respective JM109 (DE3) *E. coli* cells with the plasmid constructs encoding pupylation proteins pup and C-terminal 6xHis-tagged PafA. The additional faint bands observed in each of the linearised reactions was attributed to undigested vector DNA left over during the respective digestion reactions. This observation was evidenced by the comparatively similar migration pattern adopted by the faint bands relative to the undigested reference lanes of the same plasmid construct, a pattern typically observed for undigested plasmid DNA. Multiple band migration is attributed to the multiple conformations often adopted by undigested plasmid DNA (supercoiled and nicked), resulting in friction rate variations during migration and consequently different migration speeds to yield multiple bands of the same sample (Tweedie and Stowell, 2005).

For further confirmation, Sanger sequencing of the pET24b(+)-PafA and pET24b(+)-PupEPafA constructs and amplification of the respective PafA and PupE genes using the same primers via PCR was performed. Sequencing and PCR of both plasmids confirmed that the JM109 (DE3) *E. coli* cells were successfully transformed with the relevant plasmid construct, both yielding 100% similarity to the *Mtb* proteins PafA and pup with high confidence i.e., low E-values suggesting less likelihood of the obtained sequence similarities with the respective *Mtb* protein sequence to be due to chance. Targeted amplification of PafA and PupE genes by PCR yielded products within the expected sizes i.e., 1.6 kbp for PafA (expected size 1.4 kbp) and 204 bp for PupE (expected size 207 bp). Having confirmed JM109 (DE3) *E. coli* cells transformation with both pET24b(+)-PafA and pET24b(+)-PupEPafA plasmid constructs, a BL21 (DE3) *E. coli* cell line was transformed with pET24b(+)-PafA DNA construct isolated from the pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cell line for a comparative analysis of PafA recombinant expression in two different *E. coli* strains.

4.2 PafA recombinant expression and purification

PafA was successfully recombinantly expressed in JM109 (DE3) *E. coli* cells transformed with either the pET24b(+)-PafA or pET24b(+)-PupEPafA constructs (Figure 3.3 and Figure 3.4, respectively), detected at 50-kDa size by immunoblot. This extrapolated size correlated with the expected 51-kDa reported by both Özcelik *et al*, (2012) and Zhang *et al*, (2017) for PafA. Similarly, a predicted size of 51-kDa (pI = 6.26; isoelectric point) was returned by the Expasy Compute pI/Mw online database using the PafA sequence (accession number: NP_216613.1) sourced from the NCBI database which was also used in the alignment of the pET24b(+)-PafA

construct translated Sanger sequence. The absence of PafA in non-recombinant JM109 (DE3) *E. coli* cells in the presence and absence of IPTG indicated the absence of endogenous PafA expression by the bacterial cell line. PafA expression by the pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cells only when exposed to IPTG revealed inducible PafA overexpression in the recombinant cells, making the pET24b(+)-PafA transformed cells eligible for downstream PafA expression assays. In contrast, both the control and IPTG exposed pET24b(+)-PupEPafA recombinant JM109 (DE3) *E. coli* cells sustained PafA expression, indicating PafA basal expression by the dual expression system. This basal expression was due to the disruption of the lac operon during pET24b(+)-PupEPafA construction by Cerda-Maira *et al*, (2020), where after initially cloning PafA into the *Nde*I - *Hind*III sites of the pET24b(+) vector, PupE was cloned into the *Bgl*II - *Xba*I sites - hence disrupting the lac operator. This disruption consequently affected the lac operon repression functionality, resulting in constitutive and detectable PafA expression.

Additionally, JM109 (DE3) *E. coli* cells transformed with the pET24b(+)-PupEPafA also resulted in an observably slower growth rate compared to the pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cells, leading to overall lower cell culture densities. This suggested pET24b(+)-PupEPafA toxicity towards the JM109 (DE3) *E. coli* cell line, specifically the constitutive expression of PafA and pup proteins which appeared to disrupt the native proliferation and homeostasis of the bacterial cells. Whilst cellular toxicity by one or both encoded proteins (PafA and PupE) is a possibility (Rosano and Ceccarelli, 2014), pET24b(+)-PafA recombinant PafA overexpression in the same cell line without significant effect on cell growth suggested otherwise. Moreover, free pup is highly unstable in a cellular environment i.e., easily broken down, thus lowering the likelihood of PupE toxicity against the host cells. One hypothesis for the observed toxicity could be the possible pupylation of native *E. coli* proteins using the actively expressed PafA and PupE, this modification negatively affecting the activity of some essential proteins. Pupylation of *E. coli* proteins has already been established in a study conducted by Cerda-Maira and coworkers, reporting successful pupylation of 51 *E. coli* proteins in cells recombinantly producing PafA and PupE (Cerda-Maira *et al.*, 2011). The group additionally reported a considerable decrease in overall cell growth of PafA and pup dual expressing recombinant *E. coli* cells compared to *E. coli* cells transformed with an empty vector. Interestingly, no significant cell growth variation was observed between recombinant *E. coli* cells transformed with the dual expression constructs coding for PafA and

either active PupE or inactive PupQ. This suggested the observed decrease in cell growth to be independent from endogenous pupylation of *E. coli* proteins but a consequent of constitutively expressing two foreign proteins concurrently at a high energy cost.

Interestingly, pET24b(+)-PafA recombinant BL21 (DE3) *E. coli* cells also supported basal PafA expression unlike the JM109 *E. coli* cells transformed with the same plasmid construct. Contrary to the slow growth observed through the uncontrollable expression system of the pET24b(+)-PupEPafA recombinant JM109 (DE3) *E. coli*, the recombinant PafA expressing BL21 (DE3) cells incurred no significant change in cell proliferation relative to the non-recombinant control sample. This demonstrated PafA's non-toxic nature towards the BL21 (DE3) *E. coli* cell line, further implying a combined toxic propensity of PafA and PupE observed in pET24b(+)-PupEPafA recombinant JM109 *E. coli* cells. The difference in the *E. coli* strains, however, could also be a contributing factor to the toxicity difference observed for basal expressing recombinant *E. coli* cells. To validate the concept of a synergistic toxic effect of PafA and PupE, the BL21 (DE3) *E. coli* cell line would need to be transformed with the pET24b(+)-PupEPafA construct and proliferation monitored comparatively to non-recombinant BL21 (DE3) *E. coli* cells.

Taken together, the inducible pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cells were therefore selected for further PafA expression and downstream analysis that included purification and screening pure PafA for small-molecule binding ligands via TSA. In addition to providing a regulated recombinant expression system, the higher sensitivity of JM109 (DE3) *E. coli* cells to growth conditions- compared to the commonly recruited BL21 (DE3) *E. coli* strain for recombinant protein expression- results in a lower proliferation rate which aids in protein solubility i.e., increased protein refolding time thus less aggregation.

4.3 PafA solubilisation and purification

Before purification, recombinant PafA expression in the pET24b(+)-PafA recombinant JM109 (DE3) *E. coli* cells using the previous expression conditions was assessed and PafA found exclusively in the insoluble fraction as aggregated inclusions. This correlated with the predicted solubility analysis done via the virtual bioinformatics webtool Protein-Sol (<http://protein-sol.manchester.ac.uk>), predicting PafA to be less soluble than the average soluble *E. coli* protein (predicted scale solubility = 0.350) (Niwa *et al.*, 2009). However, the observed insoluble PafA

expression was contradictory to the work done by Festa et al., 2007 and Cerda-Maira et al., both reporting soluble PafA overexpression in pET24b(+)-PafA recombinant *E. coli* strains ER2566, EHD760, and EHD826. This observed difference was credited to the difference in host strains, the host cells being an important factor to consider in recombinant protein expression as they can affect production yields and protein properties such as solubility (Gopal and Kumar, 2013).

Observing PafA exclusive expression as inclusion aggregates, low IPTG induction concentration and low temperature expression was attempted to help facilitate PafA expression in the soluble phase. Both these conditions lower the production rate of recombinant PafA inside host cells, low temperature slowing down host cell proliferation directly and low IPTG concentration causing slow PafA induction, both of which help increase protein refolding time and consequently recombinant protein solubility. Using both these conditions individually and combined, insoluble PafA recombinant expression persisted thus prompting the recruitment of protein solubilisation measures and later refolding. This was done because- despite a variety of parameters that can be modulated and/or the use of additives- methods to facilitate soluble recombinant protein expression are seldom effective and incur additional costs. As such, mild solubilisation of inclusion aggregate proteins using denaturants such as urea and guanidine hydrochloride and subsequent downstream refolding in lieu of physicochemical modulation of expression conditions is increasingly employed. In fact, the major bottleneck faced by this method pertains to large-scale operations as opposed to lab-scale, requiring excess buffer for protein refolding (see Singh *et al.*, (2015) for a thorough review on recombinant protein isolation using denaturants).

To solubilise PafA inclusion aggregates, the ionic detergent SDS was recruited as the suitable denaturant due to its general availability, its compatibility with IMAC resin used downstream for PafA purification, and simple bulk removal via low temperature precipitation. PafA solubilisation was carried out using a protocol amended from Schlager and coworkers who showcased a low-cost, fast, and easy isolation protocol of recombinant protein from inclusion aggregates using SDS (Schlager *et al.*, 2012). After PafA inclusion aggregates isolation from most cell components i.e., lysis-centrifugation-wash, 1% (w/v) SDS and sonication successfully solubilised a substantial amount of PafA. SDS solubilisation of proteins is a well characterised phenomenon that is affected by the collaborative efforts of the detergents ionic and hydrophobic interactions with the target protein. High purity PafA was successfully recovered from the

solubilised PafA inclusion aggregate sample using IMAC purification. Using SDS at the upper limit of the resin's tolerable concentration i.e., 1% (w/v), resulted in inconsistent binding efficiency of PafA to the resin which is most likely due to the charged nature of SDS thus interfering with the positively charged cobalt ions of the resin.

Despite the apparent advantages of the SDS facilitated isolation of recombinant inclusion aggregate proteins, the methodology is not without caveats. One major concern is the strong denaturant properties of SDS, contrasting the mild denaturation conditions offered by other typically recruited denaturants to help favour correct protein refolding downstream. Despite this, protein refolding from a SDS denatured state have been reported. A study by Otzen and Oliveberg reported efficient refolding of proteins chymotrypsin inhibitor 2, S6, and lysozyme, by the rapid chelation of SDS by oligosaccharide α -cyclodextrin (Otzen and Oliveberg, 2001). Similarly, work done by Pederson and coworkers demonstrated favourable β -lactoglobulin protein refolding using the non-ionic surfactant octaethyleneglycol monododecyl ether. Moreover, proteins with tertiary structures consisting of disulphide links appear theoretically favoured by SDS facilitated recovery, the detergent being unable to break disulphide linkages thus providing a pseudo "mild denaturation" helping favour correct refolding downstream. To assess PafA refolding efficacy, a PafA enzyme assay will need to be developed and optimised, testing for PafA pupylating activity *in vitro*.

4.4 Non-specific inhibitor assay

Non-specific activity compounds present a major obstacle in drug discovery, prominently giving false positives during the drug development process by disguising themselves as potential drug candidates during the initial compound screening stage. This often leads to the waste of precious time and resources where research is dedicated to a compound that is not suitable for drug development. Therefore, the development of assays capable of detecting PAIN compounds as a preliminary screen in early drug development is ideal.

A non-specific inhibitor detecting assay was developed by amending protocols by Feng and Shoichet, (2006) and Ryan et al, (2003). Using the enzyme substrate duo HRP-TMB, a commonly recruited colorimetric assay particularly in enzyme-linked immunosorbent assays, the assay was optimised. Increasing the enzyme concentration to 100 ng/mL significantly increased spectrophotometric detection ($p < 0.0001$) to acceptable levels (Figure 3.9). High

detection was attributed to the resulting increase in the assays' coloured product formation. Increasing the enzyme concentration results in comparatively more enzyme molecules interacting with the TMB substrate causing more product formation per unit time i.e., increased reaction rate, hence the higher detection observed at appreciably higher enzyme concentrations.

The optimised assay was then recruited in the development of a non-specific inhibitor detecting assay using various compounds of known veracity. As anticipated, the two negative control compounds- caffeine and warfarin- did not significantly affect the colorimetric assay, which was used to emphasise the enzymes stability against compounds with specific targets (i.e., specific inhibitors). Caffeine is found in common beverages like coffee and is a well-known antagonist of the adenosine A_{2A} receptors resulting in a wakefulness effect (López-Cruz et al., 2018). On the other hand, warfarin is a clinical anticoagulant drug that competitively inhibits the vitamin K epoxide reductase complex 1 (Rettie and Tai, 2006). Riboflavin (vitamin B2) also observed no significant effect on the colorimetric assay despite the anticipated interference owing to its pigmented nature i.e., the strong yellow pigment interfering with the reactions blue colour product formation. This suggested the assays stability against pigmented compounds, thus showcasing the assays versatility at screening an array of chemical compounds. The positive control NaN₃ resulted in significant HRP inhibition ($p < 0.01$), yielding inhibition activity > 50% as was expected. NaN₃ inhibition of HRP activity by 60.6% was due to using a higher NaN₃ concentration of 1.5 mM than the reported K_i of 1.47 mM (Ortiz de Montellano *et al.*, 1988). NaN₃ inhibition of HRP at correlating reported concentrations showcased the enzymes viability and that it could be inhibited.

Plants commonly use a cocktail of non-specific inhibiting compounds as a defence mechanism to abrogate infectious agents (Cowan, 1999; Meier *et al.*, 2017). It was for this reason that model plant-derived compounds were selected to develop and optimise the non-specific inhibitor detecting assay. Three of these compounds i.e., EGCG, quercetin, and tannic acid resulted in potent inhibiting activity (> 95% inhibition) of the assay, thus showcasing the susceptibility of the assay to non-specific inhibitor compounds. The assay was also considerably inhibited by 8-hydroxyquinoline, inhibiting the assay by 76.1%. Inhibition of the assay was postulated to be caused by HRP sequestration from the reaction by the aggregating inhibitors, a mechanism reported in a study by McGovern and colleagues which gave insight on the mode of action that is typically employed by aggregating inhibitors (McGovern et al.,

2003). Using electron microscopy, the group demonstrated equal inhibition of two β -Lactamase mutants (one thermostable) via their association with the surface of aggregate structures formed by relevant aggregator inhibitors, consequently sequestering the two proteins indiscriminately out of solution. Similarly, Bustos and coworkers observed intestinal lipase sequestration by EGCG and quercetin via aggregation activity, each compound respectively reducing lipase in solution by 30% and 50% (Bustos *et al.*, 2020). This therefore suggested that the observed HRP inhibition was a consequence of the same promiscuous activity by the aggregator compounds.

A 48-compound set from the DiverSET CHEMBRIDGE small-molecule chemical library was selected to screen for potential PafA-binding ligands. Before screening the compound library, however, the developed non-specific inhibitor detecting assay was recruited as a preliminary screen to discern and eliminate any compounds from the library displaying non-specific inhibition activity. Assayed at 10 μ M, all 48 compounds from the DiverSET CHEMBRIDGE chemical library did not result in significant inhibition of the colorimetric assay ($p > 0.05$). Though statistically insignificant due to high data variability, the inhibiting effect that compound C11 exerted on HRP raised concerns. Despite varying, all three replicate's assays of the compound showed some degree of HRP inhibition activity, suggesting the compound to possess some kind of non-specific inhibitory activity. Suggesting no non-specific inhibition activity, all 48 compounds from the DiverSET CHEMBRIDGE chemical library were found appropriate and therefore selected for further downstream screening for potential PafA-binding ligands.

4.5 Screening for PafA-binding ligands via a TSA

To screen for potential PafA-binding compounds from the 48-set DiverSET CHEMBRIDGE compound library, a TSA was designed and optimised. Before the assay was developed, however, a control TSA was conducted using a control protein with an expected T_m range (41 - 47 $^{\circ}$ C) and a control ligand. The conducted control TSA yielded acceptable data within the expected ranges, shown with a T_m of 40.9 ± 0.5 $^{\circ}$ C. This implied that the thermal melt reactions were aptly conducted and that non-linear regression analysis using Graphpad Prism could also appropriately extrapolate T_m values from thermal melt curves generated from TSA data. Incubation of the control protein with the control ligand considerably increased the control proteins T_m by 5.7 $^{\circ}$ C. The observed shift in T_m - and the control ligand being the only variable

among the constant conditions of the control protein thermal melt reaction- indicated the control ligands binding influence on the control protein which subsequently increased the proteins thermostability. Protein thermostability resulting from ligand binding is linked to the ligands interaction with the protein coupling with the proteins unfolding equilibrium (Celej *et al.*, 2003). This was showcased in a study by González and coworkers who reported a staggering 37 °C T_m increase for streptavidin at elevated biotin ligand concentrations (González *et al.*, 1997).

Using the same cycle conditions from the control protein TSA, a PafA TSA was designed and optimised by varying PafA concentration and SYPRO Orange dye concentrations. The combination of 10 ng/μL pure PafA and 1 x SYPRO Orange dye provided the optimum thermal melt conditions for the protein, yielding the characteristic Sigmoidal thermal melt curves similarly observed in the control protein assays. Analysing the data PafA T_m was revealed at 47.8 ± 1.2 °C, falling within the T_m expected range for a protein from a mesophilic organism. Similarly, Khrapunov and Brenowitz reported an *E. coli* recombinantly expressed *Mtb* protein, pentapeptide-repeat protein, with a comparative T_m of 51.3 ± 0.3 °C (Khrapunov and Brenowitz, 2011). T_m is affected by a range of factors including protein size (see Kumar *et al.*, 2000), therefore proteins of the same organism typically have different T_m 's. The optimised PafA thermal melt reaction was subsequently recruited to screen compounds from the DiverSET CHEMBRIDGE chemical library, previously screened for non-specific inhibiting activity, for potential PafA-binding ligands. Eight compounds from the library resulted in a significant shift of PafA T_m i.e., ≥ 2 °C increase in PafA T_m and $P < 0.05$. Inducing T_m shifts > 5 °C, all 8 compounds suggested high binding affinity to PafA, compound C9 displaying the highest potential with a T_m shift of 10.1 °C.

Despite showing significant effect on PafA, however, Compound C11 was eliminated as a potential PafA-binding ligand on suspicions of being a non-specific inhibitor. Though the method of inhibition was unclear, having already yielded inconclusive result in the non-specific inhibitor assay- resulting in varying degrees of HRP inhibiting activity, and now binding PafA, Compound C11 inconsistent data gave enough grounds to eliminate the compound from further downstream analysis.

The seven significant PafA binders, were therefore selected for further downstream screening assays in the development of a novel PROTAC-like prokaryotic degrading drug. Future work

will look to screen and discern the seven compounds for PafA binders from inhibitors via a developed and optimised PafA inhibitor assay. PafA binders with no inhibition activity would then be selected to synthesise the PROTAC-like molecule.

REFERENCES

1. An, S. & Fu, L. 2018. Small-molecule PROTACs: An emerging and promising approach for the development of targeted therapy drugs. *EBioMedicine*, 36, 553-562.
2. Barandun, J., Delley, C. L. & Weber-Ban, E. 2012. The pupylation pathway and its role in *Mycobacteria*. *BMC Biology*, 10, 1-9.
3. Bertrand, M.J., Milutinovic, S., Dickson, K.M., Ho, W.C., Boudreault, A., Durkin, J., Gillard, J.W., Jaquith, J.B., Morris, S.J. & Barker, P.A. 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell*, 30, 689-700.
4. Bhaduri, U. & Merla, G. 2021. Ubiquitination, Biotech Startups, and the Future of TRIM Family Proteins: A TRIM-Endous Opportunity. *Cells*, 10, 1015.
5. Blanchard, J.S., 1996. Molecular mechanisms of drug resistance in *Mycobacterium tuberculosis*. *Annual Review of Biochemistry*, 65, 215-239.
6. Bond, M.J. & Crews, C.M. 2021. Proteolysis targeting chimeras (PROTACs) come of age: entering the third decade of targeted protein degradation. *RSC Chemical Biology*, 2, 725-742.
7. Bondeson, D. P. & Crews, C. M. 2017. Targeted protein degradation by small molecules. *Annual Review of Pharmacology and Toxicology*, 57, 107-123.
8. Bondeson, D. P., Mares, A., Smith, I. E., Ko, E., Campos, S., Miah, A. H., Mulholland, K. E., Routly, N., Buckley, D. L. & Gustafson, J. L. 2015. Catalytic *in vivo* protein knockdown by small-molecule PROTACs. *Nature Chemical Biology*, 11, 611-617.
9. Burns, K. E. & Darwin, K. H. 2010. Pupylation versus ubiquitylation: tagging for proteasome-dependent degradation. *Cellular Microbiology*, 12, 424-431.
10. Bustos, A.S., Håkansson, A., Linares-Pastén, J.A., Peñarrieta, J.M. & Nilsson, L. 2020. Interaction of quercetin and epigallocatechin gallate (EGCG) aggregates with pancreatic lipase under simplified intestinal conditions. *PloS One*, 15, 0224853.
11. Carreira, E. M., Pfaff, P., Samarasinghe, K. T. & Crews, C. M. 2019. Reversible spatiotemporal control of induced protein degradation by bistable photoPROTACs. *ACS Central Science*, 5, 1682-1690.
12. Cecchini, C., Pannilunghi, S., Tardy, S. & Scapozza, L. 2021. From conception to development: investigating PROTACs features for improved cell permeability and successful protein degradation. *Frontiers in Chemistry*, 9.

13. Celej, M.S., Montich, G.G. & Fidelio, G.D. 2003. Protein stability induced by ligand binding correlates with changes in protein flexibility. *Protein Science*, 12, 1496-1506.
14. Cerda-Maira, F. A., McAllister, F., Bode, N. J., Burns, K. E., Gygi, S. P., & Darwin, K. H. 2011. Reconstitution of the *Mycobacterium tuberculosis* pupylation pathway in *Escherichia coli*. *EMBO Reports*, 12, 863-870.
15. Ciulli, A., Zengerle, M. & Chan, K. H. 2015. Selective small molecule induced degradation of the BET bromodomain protein BRD4. *ACS Chemical Biology*, 10, 1770-1777.
16. Cowan, M.M. 1999. Plant products as antimicrobial agents. *Clinical Microbiology Reviews*, 12, 564-582.
17. Cui, H., Müller, A.U., Leibundgut, M., Tian, J., Ban, N. & Weber-Ban, E. 2021. Structures of prokaryotic ubiquitin-like protein Pup in complex with depupylase Dop reveal the mechanism of catalytic phosphate formation. *Nature Communications*, 12, 1-12.
18. Delport, A. & Hewer, R. 2019. Inducing the degradation of Disease-related proteins using heterobifunctional molecules. *Molecules*, 24(18), 3272.
19. Deng, L., Meng, T., Chen, L., Wei, W. & Wang, P. 2020. The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduction and Targeted Therapy*, 5, 1-28.
20. Dong, S., Chen, H., Zhou, Q. & Liao, N. 2021. Protein degradation control and regulation of bacterial survival and pathogenicity: the role of protein degradation systems in bacteria. *Molecular Biology Reports*, 48, 7575-7585.
21. Dragovich, P. S., Pillow, T. H., Adhikari, P., Blake, R. A., Chen, J., Del Rosario, G., Deshmukh, G., Figueroa, I., Gascoigne, K. E., Kamath, A. V. & Kaufman, S. 2020. Antibody conjugation of a chimeric BET degrader enables *in vivo* activity. *ChemMedChem*, 15, 17-25.
22. Elharar, Y., Roth, Z., Hermelin, I., Moon, A., Peretz, G., Shenkerman, Y., Vishkautzan, M., Khalaila, I. & Gur, E., 2014. Survival of *Mycobacteria* depends on proteasome-mediated amino acid recycling under nutrient limitation. *The EMBO Journal*, 33(16), 1802-1814.
23. Feng, B.Y. & Shoichet, B.K. 2006. A detergent-based assay for the detection of promiscuous inhibitors. *Nature Protocols*, 1, 550-553.

24. Festa, R.A., Pearce, M.J. & Darwin, K.H. 2007. Characterization of the proteasome accessory factor (paf) operon in *Mycobacterium tuberculosis*. *Journal of Bacteriology*, 189, 3044-3050.
25. Fischer, E.S., Böhm, K., Lydeard, J.R., Yang, H., Stadler, M.B., Cavadini, S., Nagel, J., Serluca, F., Acker, V., Lingaraju, G.M. & Tichkule, R.B., 2014. Structure of the DDB1–CRBN E3 ubiquitin ligase in complex with thalidomide. *Nature*, 512(7512), 49-53.
26. Fryland, T., Christensen, J.H., Pallesen, J., Mattheisen, M., Palmfeldt, J., Bak, M., Grove, J., Demontis, D., Blechingberg, J., Ooi, H.S. & Nyegaard, M. 2016. Identification of the BRD1 interaction network and its impact on mental disorder risk. *Genome Medicine*, 8, 1-20.
27. Galloway, C.A., Sowden, M.P. & Smith, H.C., 2003. Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques*, 34(3), pp.524-530.
28. George, A. J., Hoffiz, Y. C., Charles, A. J., Zhu, Y. & Mabb, A. M. 2018. A comprehensive atlas of E3 ubiquitin ligase mutations in neurological disorders. *Frontiers in Genetics*, 9, 29.
29. Gerry, C.J. & Schreiber, S.L. 2018. Chemical probes and drug leads from advances in synthetic planning and methodology. *Nature Reviews Drug Discovery*, 17, 333-352.
30. Gillespie, S.H., 2002. Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective. *Antimicrobial Agents and Chemotherapy*, 46(2), 267-274.
31. González, M., Bagatolli, L.A., Echabe, I., Arrondo, J.L., Argaraña, C.E., Cantor, C.R. & Fidelio, G.D. 1997. Interaction of biotin with streptavidin: thermostability and conformational changes upon binding. *Journal of Biological Chemistry*, 272, 11288-11294.
32. Gopal, G.J. & Kumar, A., 2013. Strategies for the production of recombinant protein in *Escherichia coli*. *The Protein Journal*, 32, 419-425.
33. Gopal, P. & Dick, T. 2020. Targeted protein degradation in antibacterial drug discovery? *Progress in Biophysics and Molecular Biology*, 152, 10-14.
34. Hayes, C. S. & Keiler, K. C. 2010. Beyond ribosome rescue: tmRNA and co-translational processes. *FEBS Letters*, 584, 413-419.
35. He, L., Chen, C., Gao, G., Xu, K. & Ma, Z. 2020. ARV-825-induced BRD4 protein degradation as a therapy for thyroid carcinoma. *Aging (Albany NY)*, 12, 4547.

36. Huang, X. & Dixit, V.M. 2016. Drugging the undruggables: exploring the ubiquitin system for drug development. *Cell Research*, 26, 484-498.
37. Humbard, M. A., Miranda, H. V., Lim, J.-M., Krause, D. J., Pritz, J. R., Zhou, G., Chen, S., Wells, L. & Maupin-Furlow, J. A. 2010. Ubiquitin-like small archaeal modifier proteins (SAMPs) in *Haloferax volcanii*. *Nature*, 463, 54-60
38. Ichinose, Y., Genka, K., Koike, T., Kato, H., Watanabe, Y., Mori, T., Iio, S., Sakuma, A. & Ohta, M. 2003. Randomized double-blind placebo-controlled trial of bestatin in patients with resected stage I squamous-cell lung carcinoma. *Journal of the National Cancer Institute*, 95(8), pp.605-610.
39. Ishida, T. & Ciulli, A., 2021. E3 ligase ligands for PROTACs: how they were found and how to discover new ones. *SLAS DISCOVERY: Advancing the Science of Drug Discovery*, 26(4), 484-502.
40. Ito, T. & Handa, H. 2020. Molecular mechanisms of thalidomide and its derivatives. *Proceedings of the Japan Academy, Series B*, 96, 189-203.
- 41. Janssen, G. V., Zhang, S., Merks, R., Schiesswohl, C., Chatterjee, C., Darwin, K. H. & Ova, H. 2019. Discovery and Optimization of Inhibitors for the Pup Proteasome System in *Mycobacterium tuberculosis*. *BioRxiv*, 796359.**
42. Jevšvar, S., Gaberc-Porekar, V., Fonda, I., Podobnik, B., Grdadolnik, J., & Menart, V. 2005. Production of nonclassical inclusion bodies from which correctly folded protein can be extracted. *Biotechnology Progress*, 21(2), 632-639.
43. Jiang, H.-W., Czajkowsky, D. M., Wang, T., Wang, X.-D., Wang, J.-B., Zhang, H.-N., Liu, C.-X., Wu, F.-L., He, X. & Xu, Z.-W. 2018. Identification of Serine 119 as an Effective Inhibitor Binding Site of M. tuberculosis Ubiquitin-like Protein Ligase PafA Using Purified Proteins and M. smegmatis. *EBioMedicine*, 30, 225-236.
44. Kelley, D. R. 2018. E3 ubiquitin ligases: key regulators of hormone signaling in plants. *Molecular & Cellular Proteomics*, 17, 1047-1054.
45. Khrapunov, S. & Brenowitz, M. 2011. Stability, denaturation and refolding of *Mycobacterium tuberculosis* MfpA, a DNA mimicking protein that confers antibiotic resistance. *Biophysical Chemistry*, 159, 33-40.
46. Klein, K., Kato, M., Frank-Bertoncelj, M., Kolling, C., Ciurea, A., Gay, S. & Ospelt, C. 2018. Evaluating the bromodomain protein BRD1 as a therapeutic target in rheumatoid arthritis. *Scientific Reports*, 8, 1-7.

47. Krönke, J., Udeshi, N. D., Narla, A., Grauman, P., Hurst, S. N., McConkey, M. & Ebert, B. L. 2014. Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science*, 343, 301-305.
48. Kumar, S., Tsai, C.J. & Nussinov, R. 2000. Factors enhancing protein thermostability. *Protein Engineering*, 13, 179-191.
49. Li, X. & Song, Y. 2020. Proteolysis-targeting chimera (PROTAC) for targeted protein degradation and cancer therapy. *Journal of Hematology & Oncology*, 13, 1-14.
50. Liu, J., Ma, J., Liu, Y., Xia, J., Li, Y., Wang, Z.P. & Wei, W. 2020. PROTACs: a novel strategy for cancer therapy. *Seminars in Cancer Biology*, 67, 171-179.
51. López-Cruz, L., Salamone, J.D. & Correa, M. 2018. Caffeine and selective adenosine receptor antagonists as new therapeutic tools for the motivational symptoms of depression. *Frontiers in Pharmacology*, 9, 526.
52. Loveday, M., Mzobe, Y., Pillay, Y. & Barron, P. 2019. Figures of the dead: A decade of tuberculosis mortality registrations in South Africa. *SAMJ: South African Medical Journal*, 109, 728-732.
53. Luthra, S., Rominski, A. & Sander, P. 2018. The role of antibiotic-target-modifying and antibiotic-modifying enzymes in *Mycobacterium abscessus* drug resistance. *Frontiers in Microbiology*, 9, 2179.
54. Madden, S.K., de Araujo, A.D., Gerhardt, M., Fairlie, D.P. & Mason, J.M., 2021. Taking the Myc out of cancer: toward therapeutic strategies to directly inhibit c-Myc. *Molecular Cancer*, 20(1), 1-18.
55. Maupin-Furlow, J.A. 2014. Prokaryotic ubiquitin-like protein modification. *Annual Review of Microbiology*, 68, 155-175.
56. McGovern, S.L., Helfand, B.T., Feng, B. & Shoichet, B.K. 2003. A specific mechanism of nonspecific inhibition. *Journal of Medicinal Chemistry*, 46, 4265-4272.
57. Meier, A.K., Worch, S., Böer, E., Hartmann, A., Mascher, M., Marzec, M., Scholz, U., Riechen, J., Baronian, K., Schauer, F. & Bode, R. 2017. Agdc1p—a Gallic acid decarboxylase involved in the degradation of tannic acid in the yeast *Blastobotrys* (Arxula) adeninivorans. *Frontiers in Microbiology*, 8, 1777.
58. Morreale, F.E., Kleine, S., Leodolter, J., Ovchinnikov, S., Kley, J., Kurzbauer, R., Hoi, D.M., Meinhart, A., Hartl, M., Haselbach, D. & Kaiser, M. 2021. BacPROTACs mediate targeted protein degradation in bacteria. *BioRxiv*.

59. Mu, X., Bai, L., Xu, Y., Wang, J. & Lu, H. 2020. Protein targeting chimeric molecules specific for dual bromodomain 4 (BRD4) and polo-like kinase 1 (PLK1) proteins in acute myeloid leukemia cells. *Biochemical and Biophysical Research Communications*, 521, 833-839.
60. Mullard, A. 2019. First targeted protein degrader hits the clinic. *Nature Reviews. Drug Discovery*.
61. Niwa, T., Ying, B.W., Saito, K., Jin, W., Takada, S., Ueda, T. & Taguchi, H. 2009. Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. *Proceedings of the National Academy of Sciences*, 106, 4201-4206.
62. Ocaña, A., Del Mar Noblejas-López, M., Nieto-Jimenez, C., Burgos, M., Gómez-Juárez, M., Montero, J. C., Esparís-Ogando, A., Pandiella, A. & Galán-Moya, E. M. 2019. Activity of BET-proteolysis targeting chimeric (PROTAC) compounds in triple negative breast cancer. *Journal of Experimental & Clinical Cancer Research*, 38, 1-9.
63. Ortiz de Montellano, P. R., David, S. K., Ator, M. A., & Tew, D. 1988. Mechanism-based inactivation of horseradish peroxidase by sodium azide. Formation of meso-azidoproporphyrin IX. *Biochemistry*, 27(15), 5470-5476.
64. Otzen, D.E. & Oliveberg, M. 2001. A simple way to measure protein refolding rates in water. *Journal of Molecular Biology*, 313, 479-483.
65. Özcelik, D., Barandun, J., Schmitz, N., Sutter, M., Guth, E., Damberger, F. F., Allain, F. H.-T., Ban, N. & Weber-Ban, E. 2012. Structures of Pup ligase PafA and depupylase Dop from the prokaryotic ubiquitin-like modification pathway. *Nature Communications*, 3, 1-10.
66. Palomino, J.C. & Martin, A. 2014. Drug resistance mechanisms in *Mycobacterium tuberculosis*. *Antibiotics*, 3(3), 317-340.
67. Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P. & Darwin, K. H. 2008. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science*, 322, 1104-1107.
68. Pettersson, M. & Crews, C.M. 2019. PROteolysis TArgeting Chimeras (PROTACs)—past, present and future. *Drug Discovery Today: Technologies*, 31, 15-27.
69. Petzold, G., Fischer, E.S. & Thomä, N.H. 2016. Structural basis of lenalidomide-induced CK1 α degradation by the CRL4CRBN ubiquitin ligase. *Nature*, 532, 127-130.

70. Qi, S.M., Dong, J., Xu, Z.Y., Cheng, X.D., Zhang, W.D. & Qin, J.J. 2021. PROTAC: An Effective Targeted Protein Degradation Strategy for Cancer Therapy. *Frontiers in Pharmacology*, 12, 1124.
71. Ramaswamy, S. & Musser, J.M., 1998. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. *Tubercle and Lung Disease*, 79(1), 3-29.
72. Rape, M. 2018. Ubiquitylation at the crossroads of development and disease. *Nature reviews Molecular Cell Biology*, 19, 59-70.
73. Rehman, W., Arfons, L. M. & Lazarus, H. M. 2011. The rise, fall and subsequent triumph of thalidomide: lessons learned in drug development. *Therapeutic Advances in Hematology*, 2, 291-308.
74. Rettie, A.E. & Tai, G. 2006. The pharmacogenomics of warfarin. *Molecular Interventions*, 6, 223.
75. Rosano, G.L. & Ceccarelli, E.A. 2014. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Frontiers in Microbiology*, 5, 172.
76. Ryan, A.J., Gray, N.M., Lowe, P.N. & Chung, C.W. 2003. Effect of detergent on “promiscuous” inhibitors. *Journal of Medicinal Chemistry*, 46, p3448-3451.
77. Sakamoto, K.M., Kim, K.B., Kumagai, A., Mercurio, F., Crews, C.M. & Deshaies, R.J., 2001. Protacs: Chimeric molecules that target proteins to the Skp1–Cullin–F box complex for ubiquitination and degradation. *Proceedings of the National Academy of Sciences*, 98(15), 8554-8559.
78. Schlager, B., Straessle, A., & Hafen, E. 2012. Use of anionic denaturing detergents to purify insoluble proteins after overexpression. *BMC Biotechnology*, 12(1), 1-7.
79. Schneekloth, A.R., Pucheault, M., Tae, H.S. & Crews, C.M., 2008. Targeted intracellular protein degradation induced by a small molecule: En route to chemical proteomics. *Bioorganic & Medicinal Chemistry Letters*, 18(22), 5904-5908.
80. Sekine, K., Takubo, K., Kikuchi, R., Nishimoto, M., Kitagawa, M., Abe, F., Nishikawa, K., Tsuruo, T. & Naito, M. 2008. Small molecules destabilize cIAP1 by activating auto-ubiquitylation. *Journal of Biological Chemistry*, 283, 8961-8968.
81. Singh, A., Upadhyay, V., Upadhyay, A.K., Singh, S.M. & Panda, A.K., 2015. Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, 14, 1-10.

82. Striebel, F., Hunkeler, M., Summer, H. & Weber-Ban, E. 2010. The mycobacterial Mpa-proteasome unfolds and degrades pupylated substrates by engaging Pup's N-terminus. *The EMBO Journal*, 29, 1262-1271.
83. Sun, D., Jeannot, K., Xiao, Y. & Knapp, C.W. 2019. Horizontal gene transfer mediated bacterial antibiotic resistance. *Frontiers in Microbiology*, 10, 1933.
84. Telenti, A., Imboden, P., Marchesi, F., Matter, L., Schopfer, K., Bodmer, T., Lowrie, D., Colston, M.J. & Cole, S., 1993. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *The Lancet*, 341(8846), 647-651.
85. Towbin, H., Staehelin, T., & Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences*, 76(9), 4350-4354.
86. Tweedie, J.W. & Stowell, K.M. 2005. Quantification of DNA by agarose gel electrophoresis and analysis of the topoisomers of plasmid and M13 DNA following treatment with a restriction endonuclease or DNA topoisomerase I. *Biochemistry and Molecular Biology Education*, 33, 28-33.
87. Wada, T., Asahi, T. & Sawamura, N., 2016. Nuclear cereblon modulates transcriptional activity of Ikaros and regulates its downstream target, enkephalin, in human neuroblastoma cells. *Biochemical and Biophysical Research Communications*, 477(3), 388-394.
88. Walters, B. J., Azam, A. B., Gillon, C. J., Josselyn, S. A. & Zovkic, I. B. 2016. Advanced *in vivo* use of CRISPR/Cas9 and anti-sense DNA inhibition for gene manipulation in the brain. *Frontiers in Genetics*, 6, 362.
89. Wang, Y., Jiang, X., Feng, F., Liu, W. & Sun, H. 2020. Degradation of proteins by PROTACs and other strategies. *Acta Pharmaceutica Sinica B*, 10, 207-238.
90. Wilson, B.A., Garud, N.R., Feder, A.F., Assaf, Z.J. & Pennings, P.S. 2016. The population genetics of drug resistance evolution in natural populations of viral, bacterial and eukaryotic pathogens. *Molecular Ecology*, 25, 42-66.
91. Winter, G. E., Buckley, D. L., Paulk, J., Roberts, J. M., Souza, A., Dhe-Paganon, S. & Bradner, J. E. 2015. Phthalimide conjugation as a strategy for *in vivo* target protein degradation. *Science*, 348, 1376-1381.
92. Winter, G.E., Buckley, D.L., Paulk, J., Roberts, J.M., Souza, A., Dhe-Paganon, S. & Bradner, J.E. 2015. Selective target protein degradation via phthalimide conjugation. *Science*, 348, 1376.

93. Zhang, S., Burns-Huang, K.E., Janssen, G.V., Li, H., Ovaa, H., Hedstrom, L. & Darwin, K.H. 2017. *Mycobacterium tuberculosis* proteasome accessory factor a (PafA) can transfer prokaryotic ubiquitin-like protein (pup) between substrates. *MBio*, 8, e00122-17.
94. Zhou, B., Hu, J., Xu, F., Chen, Z., Bai, L., Fernandez-Salas, E., Lin, M., Liu, L., Yang, C.-Y. & Zhao, Y. 2018. Discovery of a small-molecule degrader of bromodomain and extra-terminal (BET) proteins with picomolar cellular potencies and capable of achieving tumor regression. *Journal of Medicinal Chemistry*, 61, 462-481.

APPENDIX

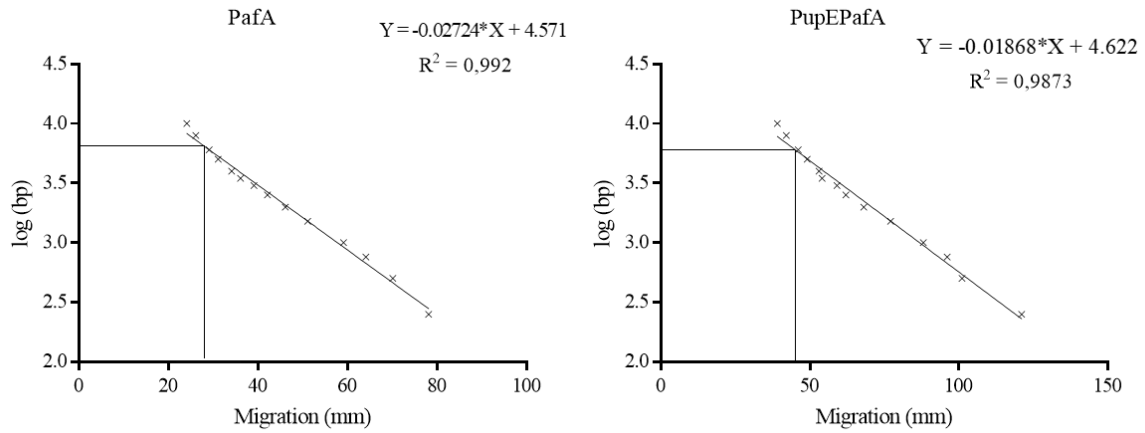


Figure A1. Vector construct size estimation. (A) pET24b(+)-PafA and (B) pET24b(+)-PupEPafA constructs were analysed on a 1% (w/v) agarose gel with a known O'GeneRuler 1 kb DNA standard (Thermo Fisher Scientific, Massachusetts, USA) and stained with 0.005% (v/v) ethidium bromide. Rf values were calculated for each DNA standard and curves plotted to extrapolate each linearised vector DNA molecular weight based on migration.

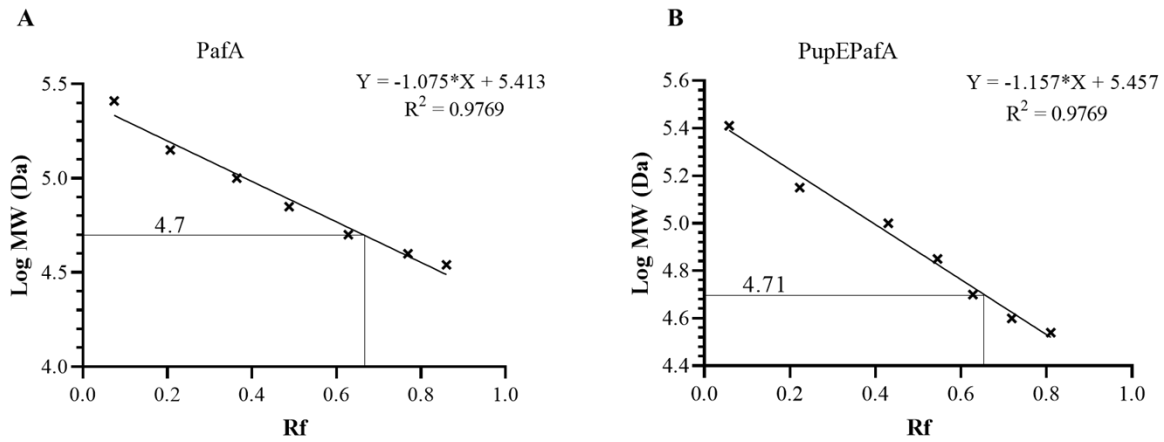


Figure A2. Recombinant PafA size estimation using known protein standards. (A) pET24b(+)-PafA and (B) pET24b(+)-PupEPafA recombinant JM109 *E. coli* cells expressing PafA were analysed on a 10% reducing SDS-PAGE with a known Spectra Multicolour Broad Range Protein standard and stained with bromophenol blue dye. Rf values were calculated for each protein standards and curves plotted to extrapolate PafA molecular weight based on migration

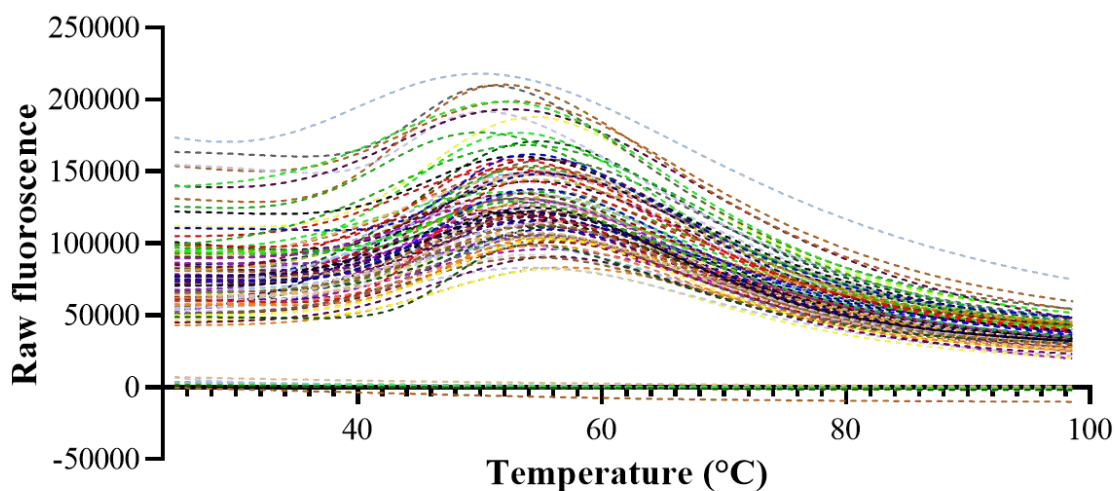


Figure A3. The effect of DiverSET CHEMBRIDGE library compounds on PafA thermal melt reaction

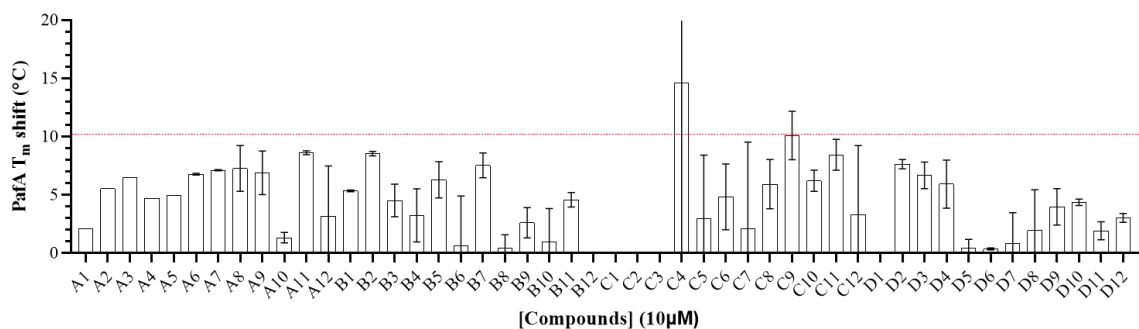
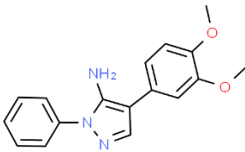
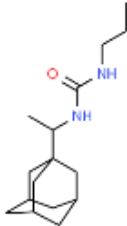
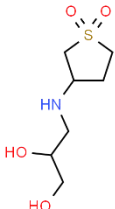
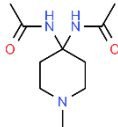
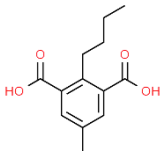
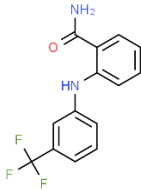


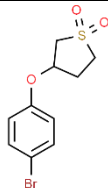
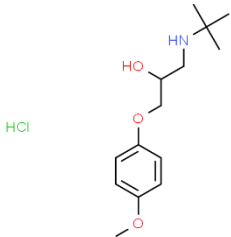
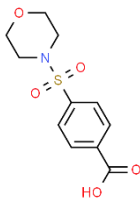
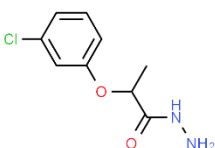
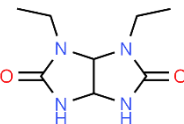
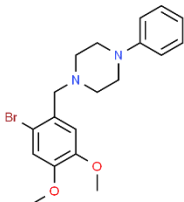
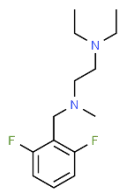
Figure A4. The calculated change in T_m of PafA in the presence of the DiverSET CHEMBRIDGE compounds. PafA (10 ng / μ L) was preincubated in PBS-MilliQ deionized water in the presence or absence (DMSO control) of 10 μ M compound for 30 min on ice. SYPRO Orange dye stock solution was subsequently added to a final 1 x concentration and the solution gently mixed, centrifuged, and PafA thermal melt assessed from 25 $^{\circ}$ C to 99 $^{\circ}$ C at 470 ± 15 nm excitation and 586 ± 10 nm emission wavelengths. The change in PafA T_m (T_m values extrapolated from the respective thermal melt reactions and taken as the EC50_1 after non-linear regression analysis) as compared to the DMSO control is shown as mean \pm standard deviation ($n = 2$).

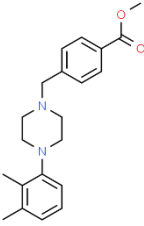
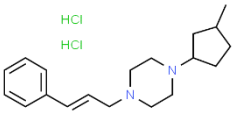
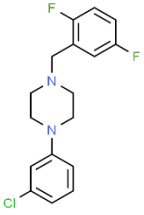
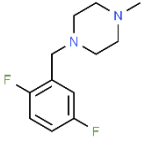
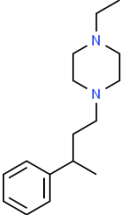
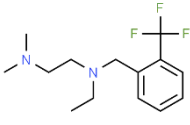
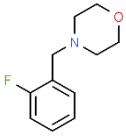
Table A1. Effect of DiverSET CHEMBRIDGE library compounds on PafA T_m

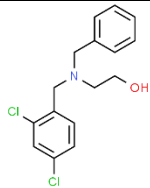
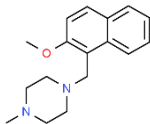
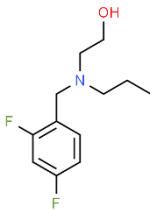
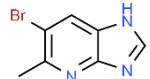
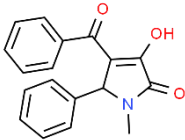
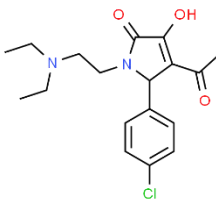
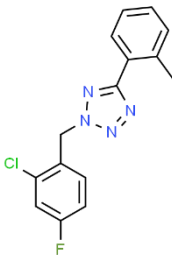
| Compound | Melting T _m (°C) | T _m shift (°C) | P-value | R-squared | Significance (<i>P</i> < 0.05) |
|----------|--------------------------------|------------------------------|---------|-----------|------------------------------------|
| A1 | 49.9 ± 3.4 | 2.1 | 0.5321 | 0.3610 | No |
| A2 | 53.0 ± 0.4 | 5.2 | 0.0436 | 0.9452 | Yes |
| A3 | 54.2 ± 1.4 | 6.4 | 0.0375 | 0.9295 | Yes |
| A4 | 52.5 ± 4.2 | 4.7 | 0.3424 | 0.6689 | No |
| A5 | 52.7 ± 6.4 | 4.9 | 0.4681 | 0.5184 | No |
| A6 | 54.5 ± 0.1 | 6.7 | 0.0782 | 0.9845 | No |
| A7 | 54.9 ± 0.1 | 7.1 | 0.0749 | 0.9860 | No |
| A8 | 55.0 ± 2.0 | 7.2 | 0.0656 | 0.9236 | No |
| A9 | 54.6 ± 1.9 | 6.8 | 0.0645 | 0.9186 | No |
| A10 | 49.1 ± 0.5 | 1.3 | 0.3406 | 0.6267 | No |
| A11 | 56.4 ± 0.2 | 8.6 | 0.0581 | 0.9899 | No |
| A12 | 50.9 ± 4.3 | 3.1 | 0.4807 | 0.4660 | No |
| B1 | 53.1 ± 0.1 | 5.3 | 0.0990 | 0.9753 | No |
| B2 | 56.3 ± 0.2 | 8.5 | 0.0568 | 0.9895 | No |
| B3 | 52.3 ± 1.4 | 4.5 | 0.0755 | 0.8615 | No |
| B4 | 51.0 ± 2.3 | 3.2 | 0.2559 | 0.6767 | No |
| B5 | 54.0 ± 1.6 | 6.2 | 0.0516 | 0.9159 | No |
| B6 | 48.4 ± 4.2 | 0.6 | 0.8640 | 0.0369 | No |
| B7 | 55.3 ± 1.1 | 7.5 | 0.0227 | 0.9571 | Yes |
| B8 | 48.1 ± 1.2 | 0.3 | 0.7684 | 0.05365 | No |
| B9 | 50.4 ± 1.3 | 2.6 | 0.1733 | 0.6865 | No |
| B10 | 48.7 ± 2.9 | 0.9 | 0.7275 | 0.1208 | No |
| B11 | 52.3 ± 0.6 | 4.5 | 0.0689 | 0.9387 | No |
| B12 | 47.6 ± 5.1 | - 0.2 | 0.9825 | 0.0007 | No |
| C1 | 35.6 ± 11.0 | - 12.2 | 0.3603 | 0.7018 | No |
| C2 | 46.3 ± 58.3 | - 1.5 | 0.9778 | 0.0012 | No |
| C3 | 7.5 ± 3.6 | 40.3 | 0.0247 | 0.9946 | Yes |
| C4 | 62.4 ± 40.10 | 14.6 | 0.6969 | 0.2096 | No |
| C5 | 50.7 ± 5.5 | 2.9 | 0.5822 | 0.3379 | No |
| C6 | 52.6 ± 2.8 | 4.8 | 0.2133 | 0.7868 | No |
| C7 | 49.9 ± 7.4 | 2.1 | 0.7538 | 0.1334 | No |
| C8 | 53.7 ± 2.1 | 5.9 | 0.1028 | 0.8828 | No |
| C9 | 57.8 ± 2.09 | 10.1 | 0.0451 | 0.9568 | Yes |
| C10 | 53.9 ± 0.9 | 6.1 | 0.0330 | 0.9483 | Yes |
| C11 | 56.2 ± 1.3 | 8.4 | 0.0224 | 0.9575 | Yes |
| C12 | 51.1 ± 5.9 | 3.3 | 0.5686 | 0.3620 | No |
| D1 | 45.71 ± 0.5 | 2.1 | 0.2124 | 0.7765 | No |
| D2 | 55.4 ± 0.40 | 7.6 | 0.0484 | 0.9835 | Yes |
| D3 | 55.4 ± 1.15 | 7.6 | 0.0296 | 0.9420 | Yes |
| D4 | 53.7 ± 2.06 | 5.9 | 0.0981 | 0.8852 | No |
| D5 | 48.2 ± 0.7 | 0.4 | 0.6951 | 0.1179 | No |
| D6 | 48.1 ± 0.1 | 0.3 | 0.7404 | 0.1553 | No |
| D7 | 48.6 ± 2.6 | 0.8 | 0.7268 | 0.1144 | No |
| D8 | 49.7 ± 3.5 | 1.9 | 0.5711 | 0.3120 | No |
| D9 | 51.7 ± 1.6 | 3.9 | 0.1121 | 0.8132 | No |
| D10 | 52.1 ± 0.3 | 4.3 | 0.1088 | 0.9587 | No |
| D11 | 49.7 ± 0.8 | 1.9 | 0.2183 | 0.6774 | No |
| D12 | 50.8 ± 0.4 | 3 | 0.1473 | 0.9063 | No |

Table A2. The portion of the DiverSET CHEMBRIDGE small molecule library available in our laboratory.

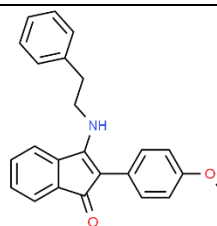
| Compound # | Name (ChemBridge ID) | Structure | Pubchem CID | Mw (g/mol) |
|------------|---|--|----------------|---------------|
| A1 | 4-(3,4-dimethoxyphenyl)-1-phenyl-1H-pyrazol-5-amine (5118317) |  | 767167 | 295.3 |
| A2 | N-[1-(1-adamantyl)ethyl]-N'-propylurea (5142981) |  | 2829891 | 264.4 |
| A3 | 3-[(1,1-dioxidotetrahydro-3-thienyl)amino]-1,2-propanediol (5155819) |  | 2831186 | 209.3 |
| A4 | N,N'-(1-methyl-4,4-piperidinediyl)diacetamide (5118841) |  | 565070 | 213.3 |
| A5 | 2-butyl-5-methylisophthalic acid (5119061) |  | 2053711 | 236.3 |
| A6 | 2-([3-(trifluoromethyl)phenyl]amino)benzamide (5144439) |  | 786469 | 280.2 |

| | | | | |
|-----|--|--|---------|-------|
| A7 | 3-(4-bromophenoxy)tetrahydrothiophene 1,1-dioxide (5155858) |  | 2831197 | 291.2 |
| A8 | N-(tert-butyl)-2-hydroxy-3-(4-methoxyphenoxy)-1-propanaminium chloride (5144520) |  | 2830264 | 289.1 |
| A9 | 4-(4-morpholinylsulfonyl)benzoic acid (5156995) |  | 559416 | 271.3 |
| A10 | 2-(3-chlorophenoxy)propanohydrazide (5189169) |  | 2832641 | 214.6 |
| A11 | 1,6-diethyltetrahydroimidazo[4,5-d]imidazole-2,5(1H,3H)-dione (5233951) |  | 789676 | 198.2 |
| A12 | 1-(2-bromo-4,5-dimethoxybenzyl)-4-phenylpiperazine (5411732) |  | 1375644 | 391.3 |
| B1 | N-(2,6-difluorobenzyl)-N',N'-diethyl-N-methyl-1,2-ethanediamine (5415800) |  | 2845331 | 256.3 |

| | | | | |
|----|---|--|----------|-------|
| B2 | methyl 4-([4-(2,3-dimethylphenyl)-1-piperazinyl]methyl)benzoate (5418608) |  | 2845408 | 338.4 |
| B3 | 1-(3-methylcyclopentyl)-4-(3-phenylprop-2-en-1-yl)piperazine dihydrochloride (5417271) |  | 45595909 | 357.4 |
| B4 | 1-(3-chlorophenyl)-4-(2,5-difluorobenzyl)piperazine (5414621) |  | 756340 | 322.8 |
| B5 | 1-(2,5-difluorobenzyl)-4-methylpiperazine (5419034) |  | 782291 | 226.3 |
| B6 | 1-ethyl-4-(3-phenylbutyl)piperazine (5430819) |  | 2846144 | 246.4 |
| B7 | N-ethyl-N',N'-dimethyl-N-[2-(trifluoromethyl)benzyl]-1,2-ethanediamine (5430906) |  | 796666 | 274.2 |
| B8 | 4-(2-fluorobenzyl)morpholine (5431331) |  | 783642 | 195.2 |

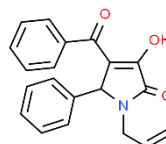
| | | | | |
|-----|--|--|---------|-------|
| B9 | 2-[benzyl(2,4-dichlorobenzyl)amino]ethanol (5431336) |  | 783643 | 309.1 |
| B10 | 1-[(2-methoxy-1-naphthyl)methyl]-4-methylpiperazine (5431935) |  | 783724 | 270.4 |
| B11 | 2-[(2,4-difluorobenzyl)(propyl)amino]ethanol (5431798) |  | 783703 | 229.3 |
| B12 | 6-bromo-5-methyl-3H-imidazo[4,5-b]pyridine (5722012) |  | 5310884 | 212.1 |
| C1 | 4-benzoyl-3-hydroxy-1-methyl-5-phenyl-1,5-dihydro-2H-pyrrol-2-one (5787093) |  | 2868343 | 293.3 |
| C2 | 4-acetyl-5-(4-chlorophenyl)-1-[2-(diethylamino)ethyl]-3-hydroxy-1,5-dihydro-2H-pyrrol-2-one (5791634) |  | 2868811 | 350.8 |
| C3 | 2-(2-chloro-4-fluorobenzyl)-5-(2-methylphenyl)-2H-tetrazole (6271753) |  | 723279 | 302.7 |

C4 2-(4-methoxyphenyl)-3-[(2-phenylethyl)amino]-1H-inden-1-one
(6194607)



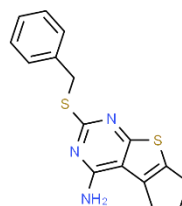
1361544 355.4

C5 1-allyl-4-benzoyl-3-hydroxy-5-phenyl-1,5-dihydro-2H-pyrrol-2-one
(6197483)



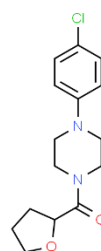
5349156 319.4

C6 2-(benzylthio)-6,7-dihydro-5H-cyclopenta[4,5]thieno[2,3-d]pyrimidin-4-amine
(6136720)



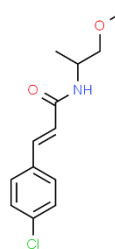
911811 313.4

C7 1-(4-chlorophenyl)-4-(tetrahydro-2-furanylcarbonyl)piperazine
(7275637)



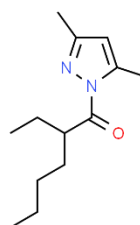
4340767 294.7

C8 3-(4-chlorophenyl)-N-(2-methoxy-1-methylethyl)acrylamide
(7364709)

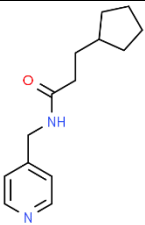
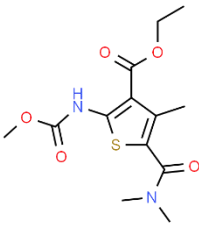
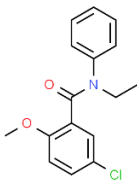
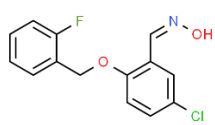
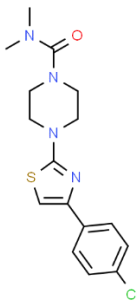
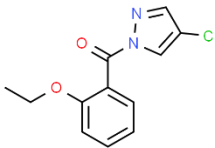


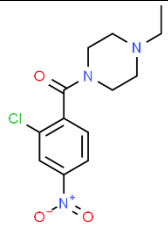
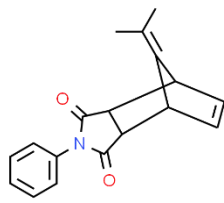
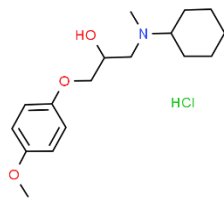
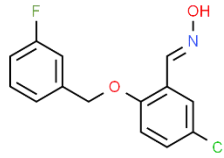
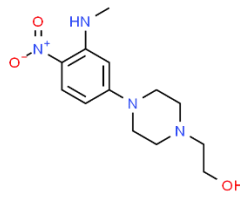
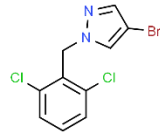
5737578 253.7

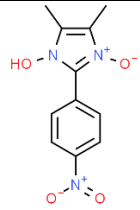
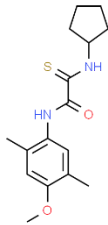
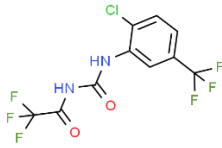
C9 1-(2-ethylhexanoyl)-3,5-dimethyl-1H-pyrazole
(7299305)



3483498 222.3

| | | | | |
|-----|---|--|---------|-------|
| C10 | 3-cyclopentyl-N-(4-pyridinylmethyl)propenamide (7293667) |  | 5087141 | 232.3 |
| C11 | ethyl 5-[(dimethylamino)carbonyl]-2-[(methoxycarbonyl)amino]-4-methyl-3-thiophenecarboxylate (7294926) |  | 887692 | 314.4 |
| C12 | 5-chloro-N-ethyl-2-methoxy-N-phenylbenzamide (6612023) |  | 670148 | 289.8 |
| D1 | 5-chloro-2-[(2-fluorobenzyl)oxy]benzaldehyde oxime (6832681) |  | 808345 | 279.0 |
| D2 | 4-[4-(4-chlorophenyl)-1,3-thiazol-2-yl]-N,N-dimethyl-1-piperazinecarboxamide (6950620) |  | 1277931 | 350.9 |
| D3 | 4-chloro-1-(2-ethoxybenzoyl)-1H-pyrazole (6660132) |  | 578476 | 250.7 |

| | | | | |
|----|--|--|----------|-------|
| D4 | 1-(2-chloro-4-nitrobenzoyl)-4-ethylpiperazine (6800944) |  | 1122354 | 297.7 |
| D5 | 10-(1-methylethylidene)-4-phenyl-4-azatricyclo[5.2.1.0~2,6~]dec-8-ene-3,5-dione (6714292) |  | 2911549 | 279.3 |
| D6 | 1-[cyclohexyl(methyl)amino]-3-(4-methoxyphenoxy)-2-propanol hydrochloride (6943696) |  | 44782365 | 329.9 |
| D7 | 5-chloro-2-[(3-fluorobenzyl)oxy]benzaldehyde oxime (6843219) |  | 797934 | 279.7 |
| D8 | 2-(4-[3-(methylamino)-4-nitrophenyl]-1-piperazinyl)ethanol (7951028) |  | 2971996 | 280.3 |
| D9 | 4-bromo-1-(2,6-dichlorobenzyl)-1H-pyrazole (7955822) |  | 1249956 | 306.0 |

| | | | | |
|-----|---|--|---------|-------|
| D10 | 4,5-dimethyl-2-(4-nitrophenyl)-1H-imidazol-1-ol 3-oxide (7959564) |  | 2057166 | 249.2 |
| D11 | 2-(cyclopentylamino)-N-(4-methoxy-2,5-dimethylphenyl)-2-thioxoacetamide (7952968) |  | 2972540 | 306.4 |
| D12 | N-([2-chloro-5-(trifluoromethyl)phenyl]amino)carbonyl-2,2,2-trifluoroacetamide (7953011) |  | 1630737 | 334.6 |